



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 98597

TO: Phillip Gambel
Location: 8b03 / 9e12
Monday, July 14, 2003
Art Unit: 1644
Phone: 308-3997
Serial Number: 09 / 825580

From: Jan Delaval
Location: Biotech-Chem Library
CM1-1E07
Phone: 308-4498

jan.delaval@uspto.gov

Search Notes

09 / 825580

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

852 280

6854Y
Delaval, Jan

From: Gambel, Phillip
Sent: Sunday, July 13, 2003 9:36 AM
To: Delaval, Jan
Subject: 09 / 825580 eppihimer amd

jan

please perform a sequence and a sequence interference search for

09 / 825,580 (eppihimer amd)

SEQ ID NO: 2

thanx

phillip gambel
art unit 1644
308-3997

1644 mailbox 9e12

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 An Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Jan Delaval
 Reference Librarian
 Biotechnology & Chemical Library
 CM1 1E07-703-308-4498
 jan.delaval@uspto.gov

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Ja</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>4498</u>	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel Orbit _____
Date searcher Proposed: <u>7/14/03</u>	Bibliographic _____	On Line _____
Date Implemented: <u>7/14/03</u>	Litigation _____	News News _____
Searcher Prep & Review Time: _____	Fulltext: _____	Sequence Systems <input checked="" type="checkbox"/>
Client Prep Time: <u>10</u>	Patent Family _____	WWW Internet _____
Indexing Time: <u>10</u>	Other _____	Other Specialty _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: July 14, 2003, 06:10:58 ; Search time 41 Seconds
(without alignments)
942.586 Million cell updates/sec

Title: US-09-825-580-2
Perfect score: 2030
Sequence: 1 MFPLQLLLILLGPGNSLQL.....TPEPREDDGLTLHSLFP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2015	99.3	412	A57468	P-selectin glycoprotein
2	288.5	14.2	2	GSFF3	salivary glue prot
3	283	13.9	867	T45463	membrane glycoprotein
4	271.5	13.4	866	T45462	membrane glycoprotein
5	256	12.6	750	T42614	probable envelope
6	250.5	12.3	797	1 VGBEX1	glycoprotein X pre
7	226.5	11.2	801	T29018	hypothetical protein
8	210.5	10.4	825	T29634	hypothetical protein
9	205.5	10.1	1151	T18535	high molecular mass
10	204.5	10.1	770	T22808	hypothetical protein
11	198	9.8	3191	T22945	hypothetical protein
12	194.5	9.6	507	T44768	antifreeze glycoprotein
13	194	9.6	1161	S57180	probable membrane
14	193.5	9.5	796	T21460	hypothetical protein
15	193	9.5	658	T41309	hypothetical protein
16	190	9.4	2761	T21064	hypothetical protein
17	189.5	9.3	3570	T25025	mucin MUC5B, trachea
18	188.5	9.3	1229	T25697	hypothetical protein
19	188.5	9.3	1829	T24583	hypothetical protein
20	187	9.2	1367	1 S48478	glucan 1,4-alpha-g
21	186.5	9.2	3020	T43932	mucin 2 precursor,
22	186	9.2	402	E86185	hypothetical protein
23	184	9.1	662	T45155	mucin FIM-C.1 - Af
24	183	9.0	1777	T34369	hypothetical protein
25	182.5	9.0	851	T22696	hypothetical protein
26	182.5	9.0	1832	T31113	mucin-like glycoprotein
27	181	8.9	235	PC2022	mucin like protein
28	180.5	8.9	400	A28172	spasmodic precursor
29	180.5	8.9	839	F75518	hypothetical protein

30	180	8.9	3507	2	T34513	hypothetical protein
31	179.5	8.8	328	2	S01359	salivary glue protein
32	178	8.8	292	2	S24169	mucin - rat
33	176.5	8.7	279	2	S53363	mucin 5AC (clone J
34	176.5	8.7	699	2	C43674	US4 protein - huma
35	175	8.6	846	2	T21700	hypothetical protein
36	174.5	8.6	232	2	A60095	larval glue protein
37	173.5	8.5	294	2	A37232	mucin, tracheal (A
38	173.5	8.5	860	2	JC4566	chitinase (EC 3.2.
39	173	8.5	216	2	I51920	mucin - rhesus mac
40	173	8.5	2225	2	T26063	hypothetical protein
41	171.5	8.4	792	2	S70305	hypothetical protein
42	171	8.4	648	2	PC4395	mucin 3 - human (f
43	171	8.4	5376	2	T42215	zonadhesin - mouse
44	170.5	8.4	354	2	T46740	microfilament shea
45	170	8.4	246	2	PC4397	mucin 3 T10 - huma

ALIGNMENTS

RESULT 1

A57468
P-selectin glycoprotein ligand PSGL-1 precursor, long splice form - human
N:contains: P-selectin glycoprotein ligand PSGL-1, long splice form; P-selectin glycoprotein
C:Species: Homo sapiens (man)
C>Date: 08-Dec-1995 #sequence revision 08-Dec-1995 #text_change 08-Oct-1999
C:Accession: A57468; A49267; A54711
R:Veldman, G.M.; Bean, K.M.; Cumming, D.A.; Eddy, R.L.; Sait, S.N.J.; Shows, T.B.
J. Biol. Chem. 270, 16470-16475, 1995
A>Title: Genomic organization and chromosomal localization of the gene encoding human P-
A:Reference number: A57468; MUID:95332364; PMID:7541799
A:Accession: A57468
A:Molecule type: DNA
A:Residues: 1-412 <VEL>
A:Cross-References: GB:U5955
R:Sako, D.; Chang, X.J.; Barone, K.M.; Vachino, G.; White, H.M.; Shaw, G.; Veldman, G.M.
Cell 75, 1179-1186, 1993
A>Title: Expression cloning of a functional glycoprotein ligand for P-selectin.
A:Reference number: A49267; MUID:94084793; PMID:7505206
A:Accession: A49267
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-134,145-412 <SAK>
A:Cross-References: GB:U02297; NID:G435416; PIDN:AAC50061.1; PID:G435417
R:Moore, K.L.; Eaton, S.F.; Lyons, D.E.; Lichenstein, H.S.; Cummings, R.D.; McEver, R.P.
J. Biol. Chem. 269, 23318-23327, 1994
A>Title: The P-selectin glycoprotein ligand from human neutrophils displays sialylated,
A:Reference number: A54711; MUID:94365038; PMID:7521878
A:Accession: A54711
A:Molecule type: protein
A:Residues: 350-355;390-391,'K',393-396 <MOO>
C:Genetics: GDB:SEUPLG
A:Gene: GDB:SEUPLG
A:Cross-References: GDB:273698; OMIM:600738
A:Map position: 12q24-12q24
A:Introns: #status absent
C:Keywords: disulfide bond; glycoprotein; polymorphism; tandem repeat; transmembrane pro
F1-412/Product; P-selectin glycoprotein ligand PSGL-1 precursor, long splice form #stat
F1-134,145-412/Product; P-selectin glycoprotein ligand PSGL-1 precursor, short splice f
F1-18/DNA: signal sequence #status predicted <SIG>
F1-319/DNA: extracellular #status predicted <EXT>
F117-261/Region: 10-residue repeats (A-T/M-E-A-Q-T-T-X-P/L-A/T)
F320-343/DNA: transmembrane #status predicted <TMW>

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 2.1e-119;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 1 MFPLQLLLILLGPGNSLQLWDWDAEAKALGPLLARDROATEYELDYDLPETEP 60
DB 1 MFPLQLLLILLGPGNSLQLWDWDAEAKALGPLLARDROATEYELDYDLPETEP 60

Qy	61	EMLRNSTDTPLTGTGCTPESTIVTEPAARRSGLDAGGAVTETLTTELAMNGNLSTDSAAME	120
Db	61	EMLRNSTDTPLTGTGCTPESTIVTEPAARRSGLDAGGAVTETLTTELAMNGNLSTDSAAME	120
Qy	121	IQTTPAA-----TEAQTPLAAATEAQTTRLTATAEAQTTPLAATEAQTTPPAATE	170
Db	121	IQTTPAAATEAQTTPVTEAQTTPLAATEAQTTRLTATAEAQTTPLAATEAQTTPPAATE	180
Qy	171	AQTTPQTGLEAQTTPAAAMEAQTTPAAAMEAQTTPAAEAQTTPAAEAQTTPAATE	230
Db	181	AQTTPQTGLEAQTTPAAAMEAQTTPAAAMEAQTTPAAEAQTTPAAEAQTTPAATE	240
Qy	231	AQTTPQTATEAQTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSVTHKGIPMAA	290
Db	241	AQTTPQTATEAQTTPLAAMEALSTEPSATEALSMEPTTKRGLFIPFSVSVTHKGIPMAA	300
Qy	291	SNLSVNVYPGADPHLSVKQCLLAILILALVATIEFVCTTVLAVRLSRKGHWYVPRNYSY	350
Db	301	SNLSVNVYPGADPHLSVKQCLLAILILALVATIEFVCTTVLAVRLSRKGHWYVPRNYSY	360
Qy	351	EMVCISLLPDGEGFSGATANGGLSKAKSPGLTPEPREDEGDDTLHSLP	402
Db	361	EMVCISLLPDGEGFSGATANGGLSKAKSPGLTPEPREDEGDDTLHSLP	412

RESULT 2

GSPF3
 salivary glue protein sgs-3 - fruit fly (*Drosophila melanogaster*)
 C/Species: *Drosophila melanogaster*
 C/Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 16-Jul-1999
 C/Accession: A03329
 R/Garfinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M.
 J. Mol. Biol. 168, 765-789, 1983
 A/Title: DNA sequences, gene regulation and modular protein evolution in the *Drosophila*
 A/Reference number: A92904; MUID:83294545; PMID:6411930
 A/Accession: A03329
 A/Molecule type: DNA
 A/Residues: 1-307 <GAR>
 A/Cross-references: GB:X01918; NID:g8581; PIDN:CAA35994.1; PID:g603989
 C/Comment: This protein is produced by third-instar larvae.
 C/Genetics:
 A/Gene: sgs-3
 A/Cross-References: FlyBase:FBgn0003373
 A/Map position: 3L (68C)
 A/Introns: 10/1
 C/Superfamily: salivary glue protein
 C/Keywords: salivary gland; tandem repeat

	Query Match	14.2%;	Score 288.5;	DB 1;	Length 307;
	Best Local Similarity	30.8%;	Pred. No. 3.2e-11;		
	Matches	69;	Conservative 32;	Mismatches 112;	Indels 11; Gaps 2
QY	55	PETEPEMLRNSTDITPLTGCTPSTETVEAA-----RRSTGLDAGGAVTELTTTEL	106		
Db	37	PRITQPCCVTTTTTTTTTTCAPTQQSITQPCTTSKPTTKPTQTTLQLP---CITPTTTKA	93		
QY	107	ANNGNLSDSAAMEIQTTOPAAATEAQTTPPLAAATEAQTRLTATEAQTTPLAAEAQTTP	166		
Db	94	TTTKPTTTKATTKATTAKTTTKQTTLQLPCLTPTTTTKQTTLQLPCTTPTTTTKPTTKP	153		
QY	167	AATEAQTTQPTGLEAQTPAAPAAEAQTTPAAPAAEAQTTPPAAMEAQTTQTNAMEAQTTAP	226		
Db	154	TTTKPTTTKPTTKTTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKP	213		
QY	227	EATEAQTTQPTAEAQTTPLAAMEALSTEPSEALSMEPTTKR	270		
Db	214	TTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKP	257		

RESULT 3

membrane glycoprotein [imported] - equine herpesvirus 1
C; Species: equine herpesvirus 1
T45463

C; Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 17-Mar-2000
C; Accession: T45463
R; Kiriisawa, R.; Kobayashi, T.; Kawakami, Y.; Iwai, H.
J. Equine Sci. 7, 79-87, 1996
A; Title: Nucleotide sequences of open reading frames 1, 24 and 71 of an att
A; Reference number: Z22973
A; Accession: T45463
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: DNA
A; Residues: 1-867 <KIR>
A; Cross-references: EMBL:D88734; PIDN:BA20038.1
A; Experimental source: isolate 3f clone; strain BK343
C; Genetics:
A; Note: ORF71
C; Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein X

	Query Match	13.9%; Score 283; DB 2; Length 867;
	Best Local Similarity	31.2%; Pred. No. 2.3e-10;
	Matches	103; Conservative 23; Mismatches 174; Indels 30; Gaps 7
Qy	67	TDTTPLTGPGPESTTVPEARRRSTGLDAGGAVTETLTTELANMGNLSDTSAMEIQTTP 126
Dd	158	TTTTPTSTTTTATTVTPTTASITDDTTTAATTAATTAATTAATTAATTAATTAATTA 217
Qy	127	AATEAQOTPLAAETAQOTRLTATAEAQTTLPLAATEAQOTPPPAAETAQOTQPTGLEAQTTAP 186
Dd	218	ATTAAATTTAAATTAATTAATTAATTAATTAATTAATSSATTAATTAATTAATTAATTA 277
Qy	187	AMBEAQOTTA PAAMBAEQOTPP PAAMBAEQOTTQTMTAMB AQOTTAPEATEAQOTTOPTATEAQTTPL 246
Dd	278	TTTAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTGAATSGS 337
Qy	247	AMNEALSTE -PSATEALSMEPTTKXGLFI PFSVSVTHKGI PMAA --SNLSVNYPVGAP 302
Dd	338	TSTTGASTSPSASTATSATPTS-----TSTSAAATTTSTPPTSAATSAESTTEAPTSTP 392
Qy	303	--DHISVKQCILAILIALVATI PFVCTVLVALRLSRKHGMYPVRNYSPTEMVCISSLLP 360
Dd	393	TTDDTTTSE-----ATTATTGPSTTVSASTISA-----TTTATFTTESHTSP 434
Qy	361	DGGEGPSATANGGLSKAKSGLPATPEPREDR 390
Dd	435	DSSTGSTSTAEPSSFTFLTPS-TATPSTDQ 463

RESULT 4

T45462
 membrane glycoprotein [imported] - equine herpesvirus 1
 C/Species: equine herpesvirus 1
 C/Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jun-2000
 C/Accession: T45462
 R/Kirisawa, R.; Kobayashi, T.; Kawakami, Y.; Iwai, H.
 J. Equine Sci. 7, 79-87, 1996
 A/Title: Nucleotide sequences of open reading frames 1, 24 and 71 of an att
 A/Reference number: Z22973
 A/Accession: T45462
 A/Status: preliminary; translated from GB/EMBL/DDBJ
 A/Molecule type: DNA
 A/Residues: 1-866 <KIR>
 A/Cross-references: EMBL:D88733; PIDN:BA20037.1
 A/Experimental source: strain HH1
 C/Genetics:
 A/Note: ORF71
 C/Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glyco
 P:558-866/Domain: equine herpesvirus 1 glycoprotein homolog <EHG>

	Query Match	13.4%	Score 271.5	DB 2	Length 866
	Best Local Similarity	30.9%	Pred. No. 1.2e-09		
	Matches 102	Conservative	23	Mismatches 170	Indels 35
					Gaps 8
Qy	67	TDTPPTGPGTPESTTVEPAARRSTGADGAVTETLTTELANMGNLSTDSAAIMEIQTTP	126		
Db	162	TTTTPTSTTTTTATTTPVTASTTDTTDTAAATTAATTTAA-----TTTAATTTAAATTTA	216		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using SW model

Run on: July 14, 2003, 06:10:58 ; Search time 23 Seconds

(without alignments)
724.934 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030
Sequence: 1 MPEQLLLILILGPGNSLQL.....TPPREDRREGDGLTHSLFLP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2015	99.3	412	1 SEPL_HUMAN	Q14442 homo sapien
2	825.5	40.7	397	1 SEPL_MOUSE	Q62170 mus musculi
3	288.5	14.2	307	1 SG63_DROME	P02840 drosophila
4	250.5	12.3	797	1 VGIX_HSVB	P28968 equine herp
5	196.5	9.7	5179	1 MUC2_HUMAN	Q02817 homo sapien
6	194	9.6	1161	1 DANA_YEAST	P47179 saccharomyc
7	193.5	9.5	3178	1 YS69_CAEEL	Q09624 caenorabd
8	187	9.2	1367	1 AMYH_YEAST	P08640 xenopus lae
9	184	9.1	662	1 MUC1_XENLA	Q05049 xenopus lae
10	180.5	8.9	400	1 MUA1_XENLA	P10667 xenopus lae
11	179.5	8.8	328	1 SG63_DROER	P13730 drosophila
12	176.5	8.7	699	1 VGIG_HSV2H	P13290 herpes simp
13	173.5	8.5	860	1 CH12_COCIM	P54197 coccidioid
14	171	8.4	5376	1 ZAN_MOUSE	O88799 mus musculi
15	169.5	8.3	1802	1 HKR1_YEAST	P41809 saccharomyc
16	168	8.3	1637	1 MRSP_STPAU	P80544 staphylococ
17	167.5	8.3	886	1 VG93_EBVA6	Q07284 epstein-bar
18	167	8.2	405	1 SDG3_CHICK	P26261 gallus gall
19	167	8.2	605	1 YH68_YEAST	P38739 saccharomyc
20	166	8.2	528	1 POD4_HUMAN	O00592 homo sapien
21	164.5	8.1	2109	1 PGCA_CHICK	P07898 gallus gall
22	164	8.1	338	1 SEB1_YEAST	O01589 saccharomyc
23	163.5	8.1	1045	1 GUNB_CELFI	P26225 cellulomona
24	163	8.0	567	1 CH13_CANAL	P40954 candida alb
25	162.5	8.0	450	1 CY1_PARDE	P13627 paracoccid
26	162.5	8.0	670	1 VG50_HSV1	Q00130 ictaluriid h
27	161	7.9	676	1 MUC1_MESAU	Q60528 mesocricetu
28	160.5	7.9	442	1 SDG3_RAT	P33671 rattus norv
29	160.5	7.9	1199	1 N121_RAT	P52591 rattus norv
30	159.5	7.9	442	1 SDG3_MOUSE	O64519 mus musculi
31	159.5	7.9	907	1 VG93_EBV	P03200 epstein-bar
32	159	7.8	1758	1 Y1R7_YEAST	P40634 saccharomyc
33	159	7.8	1758	1 YW6_YEAST	P40889 saccharomyc

ALIGNMENTS

RESULT 1
ID SEPL_HUMAN STANDARD, PRT, 412 AA.
AC Q14442; Q12775;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P-selectin glycoprotein ligand 1 precursor (PSGL-1) (Selectin P
DE ligand) (CD162 antigen).
GN SELPG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95332364; PubMed=7541799;
RA Veldman G.M., Bean K.M., Cumming D.A., Eddy R.L., Sait S.N.J.,
RA Shows T.B.;
RT "Genomic organization and chromosomal localization of the gene
RT encoding human P-selectin glycoprotein ligand.";
RL J. Biol. Chem. 270:16470-16475(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94084793; PubMed=7505206;
RA Sako D., Chang X.J., Barone K.M., Vachino G., White H.M., Shaw G.,
RA Veldman G.M., Bean K.M., Ahern T.J., Furtie B., Cumming D.A.,
RA Larsen G.R.;
RT "Expression cloning of a functional glycoprotein ligand for
RT P-selectin.";
RL Cell 75:1179-1186(1993).
RN [3]
RP SULFATION.
RX MEDLINE=96028106; PubMed=7585949;
RA Sako D., Camesa K.M., Barone K.M., Camphausen R.T., Cumming D.A.,
RA Shaw G.D.;
RT "A sulfated peptide segment at the amino terminus of PSGL-1 is
RT critical for P-selectin binding.";
RL Cell 83:323-331(1995).
RN [4]
RP SULFATION.
RX MEDLINE=96028107; PubMed=7585950;
RA Pouyau T., Seed B.;
RT "PSGL-1 recognition of P-selectin is controlled by a tyrosine
RT sulfation consensus at the PSGL-1 amino terminus.";
RL Cell 83:333-343(1995).
RN [5]
RP SULFATION.
RX MEDLINE=96032682; PubMed=7559387;
RA Wilkins P.P., Moore K.L., McEver R.P., Cummins R.D.;
RT "Tyrosine sulfation of P-selectin glycoprotein ligand-1 is required
RT for high affinity binding to P-selectin.";
RL J. Biol. Chem. 270:22677-22680(1995).
CC -1- FUNCTION: BINDS TO P-, E- AND L-SELECTINS. THE CALCIUM-DEPENDENT
HIGH AFFINITY INTERACTION WITH P-SELECTIN MEDIATES THE TETHERING

```

CC AND ROLLING OF NEUTROPHILS AND T-LYMPHOCYTES ON ENDOTHELIAL CELLS.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED ON NEUTROPHILS, MONOCYTES AND MOST
CC LYMPHOCYTES.
CC -1- PTM: HEAVILY O-GLYCOSYLATED. ALSO N-GLYCOSYLATED.
CC -1- PTM: SULFATION IN THE N-TERMINAL REGION; SULFATION IS NECESSARY
CC FOR P-SELECTIN BINDING.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD162 entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd162.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U25956; AAA74577.1; -.
CC DR EMBL; U02297; AAC50061.1; -.
CC DR GLYCOSULEDB; 014242; -.
CC DR Genew; HGNC:10722; SELPLG.
CC DR MIM; 600738; -.
CC KM Cell adhesion; Glycoprotein; Transmembrane; signal; Repeat;
CC Polymorphism; Sulfation.
CC FT SIGNAL 1 17 POTENTIAL.
CC FT PROPEP 18 41
CC FT CHAIN 42 412 P-SELECTIN GLYCOPROTEIN LIGAND 1.
CC FT DOMAIN 18 320 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 321 341 POTENTIAL.
CC FT DOMAIN 342 412 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT DOMAIN 122 261 12 X 10 AA TANDEM REPEATS.
CC FT REPEAT 122 131 1.
CC FT REPEAT 132 141 2.
CC FT REPEAT 142 151 3.
CC FT REPEAT 162 171 4.
CC FT REPEAT 182 191 5.
CC FT REPEAT 192 201 6.
CC FT REPEAT 202 211 7.
CC FT REPEAT 212 221 8.
CC FT REPEAT 222 231 9.
CC FT REPEAT 232 241 10.
CC FT REPEAT 242 251 11.
CC FT REPEAT 252 261 12.
CC FT MOD RES 46 46 SULFATION (POTENTIAL).
CC FT MOD RES 48 48 SULFATION (POTENTIAL).
CC FT MOD RES 51 51 SULFATION (POTENTIAL).
CC FT VARIANT 132 141 MISSING (IN SHORT FORM; NOT AN
CC ALTERNATIVE SPLICING).
CC /ftid=VAR_005611.
CC SQ SEQUENCE 412 AA; 43201 MW; A92A2A902DC93963A CRC64;

```

Query Match 99.3%; Score 2015; DB 1; Length 412;
 Best Local Similarity 97.6%; Pred. No. 9.4e-114;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```

QY 1 MPQLLLLLLLPGNSLQMDTWADAEKALGPLLARDRRQATEVEYLDYDELPETEP 60
Db 1 MPQLLLLLLLPGNSLQMDTWADAEKALGPLLARDRRQATEVEYLDYDELPETEP 60
QY 61 EMKRNSTDTTPLPGPGSTETVEPAARRSTGLDAGAVTETLTELAMNGNSTDSAME 120
Db 61 EMKRNSTDTTPLPGPGSTETVEPAARRSTGLDAGAVTETLTELAMNGNSTDSAME 120
QY 121 IOTTOPAA-----TEAOTPLAATAAOTPLRLATTAOTPLAATAAOTPPAAE 170
Db 121 IOTTOPAAATAAOTTPPTTAOTPLAATAAOTPLRLATTAOTPLAATAAOTPPAAE 180
QY 121 AOTTOPGLEAOTTAAPAAAEAOTTAAPAAAEAOTTPPAAMEAOTTOTTAAMEAOTTADEATE 230

```

```

Db 181 AOTTOPGLEAOTTAAPAAAEAOTTAAPAAAEAOTTPPAAMEAOTTOTTAAMEAOTTADEATE 240
QY 231 AOTTOPAAATAAOTTPPTTAOTPLAATAAOTPLRLATTAOTPLAATAAOTPPAAE 290
Db 241 AOTTOPAAATAAOTTPPTTAOTPLAATAAOTPLRLATTAOTPLAATAAOTPPAAE 300
QY 291 SNLSVAVPYGADPDHISVKOCLAILLALVATIFEVCTVLAVALSRKGMYPVNYSP 350
Db 301 SNLSVAVPYGADPDHISVKOCLAILLALVATIFEVCTVLAVALSRKGMYPVNYSP 360
QY 351 EWCICSLIPDGEGPSATANGSLSKANSPGLTPEPRDREDDTLHSFLP 402
Db 361 EWCICSLIPDGEGPSATANGSLSKANSPGLTPEPRDREDDTLHSFLP 412

```

RESULT 2
 SEPL_MOUSE
 ID SEPL_MOUSE STANDARD; PRT; 397 AA.
 AC 062170;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P-selectin glycoprotein ligand 1 precursor (PSGL-1) (selectin P ligand).
 GN SELPLG OR SELPL OR SELPL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=96320265; PubMed=8639776;
 RA Yang J., Galipeau J., Kozak C., Furie B.C., Furie B.;
 RT "Mouse P-selectin glycoprotein ligand-1: molecular cloning,
 RT chromosomal localization, and expression of a functional P-selectin
 RT receptor";
 RL Blood 87:4176-4186(1996).
 CC -1- FUNCTION: BINDS TO P-, E- AND L-SELECTINS. THE CALCIUM-DEPENDENT
 CC HIGH AFFINITY INTERACTION WITH P-SELECTIN MEDIATES THE TETHERING
 CC AND ROLLING OF NEUTROPHILS AND T-LYMPHOCYTES ON ENDOTHELIAL CELLS.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -1- PTM: HEAVILY O-GLYCOSYLATED. ALSO N-GLYCOSYLATED (BY SIMILARITY).
 CC -1- PTM: SULFATED IN THE N-TERMINAL REGION; SULFATION IS NECESSARY
 CC FOR P-SELECTIN BINDING (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X91144; CAA62583.1; -.
 CC DR MGD; MGI:106689; Selpl.
 CC KM Cell adhesion; Glycoprotein; Transmembrane; signal; Repeat;
 CC Sulfation.
 CC FT SIGNAL 1 17 POTENTIAL.
 CC FT PROPEP 18 41 BY SIMILARITY.
 CC FT CHAIN 42 397 P-SELECTIN GLYCOPROTEIN LIGAND 1.
 CC FT DOMAIN 18 307 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 308 328 POTENTIAL.
 CC FT DOMAIN 329 397 CYTOPLASMIC (POTENTIAL).
 CC FT MOD RES 54 54 SULFATION (POTENTIAL).
 CC FT MOD RES 56 56 SULFATION (POTENTIAL).
 CC FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT DOMAIN 126 225 10 X 10 AA TANDEM REPEATS.
 CC FT REPEAT 126 135 1.
 CC FT REPEAT 136 145 2.

```

FT REPEAT 146 155 3.
FT REPEAT 156 165 4.
FT REPEAT 166 175 5.
FT REPEAT 176 185 6.
FT REPEAT 186 195 7.
FT REPEAT 196 205 8.
FT REPEAT 206 215 9.
FT REPEAT 216 225 10.
SQ SEQUENCE 397 AA; 41781 MW; 512B645FFB21E4B2 CRC64;

Query Match 40.7%; Score 825.5; DB 1; Length 397;
Best Local Similarity 44.7%; Pred. No. 9.6e-43;
Matches 194; Conservative 51; Mismatches 120; Indels 69; Gaps 8;

QY 1 MPELQLLLILILPGNSLQJLMDTWADKAKALGPLLARDRRQAT-EVEYLVDYDLPETEP 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MPEPLVLLITLIGPNSLQLODPWGHETKAPGVHLRRRQVGGDDDEDDPYTNTDP 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 60 PEMLNSTDTTLP-----TGPGTPESTVEPPAARSTGLDAGAVTELT 104
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 PELKRVNTVVAHPPELPTTVVWLERDSTAGTSERATEKIAITTDPAAGTGTA----- 115
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 105 ELANNGNLSLSDSAMBIQTTOPATATATOTPLATATATOTPLATATATOTPLATATATOT 164
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 116 ---VGMLSLSDS-----TQMSLISVEVQPSSTEVES 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 165 PPAATEAQTOTPTGLEAQTATAPAAAEQTATAPAAAEQTOTPTTAMEAQT 224
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 146 QAPAPMAEISQAPAPMAEISQAPAPMAEISQAPAPMAEISQAPAPMAEIS 205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 225 APBATEAQTOTPTATATOTPL--AAMELSTEPSATALSMEPTTKGKLFIPESVS-- 279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 206 KAPAPTAETSKAPAPTAETSKAPAPTAETSKAPAPTAETSKAPAPTAETSK 265
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 280 ----SVTH-----KGIMMAASNLNVNYPVGAPEHISVKQCLLAILLALVATTEFVCT 328
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 266 FLGAPSVTHLPDGLKGLLVTPGNSPAPPLPGSSDLIPKQCLLILILASLALVFLVCT 325
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 329 VVLAVALSRKGMYPVRYNSPTFEMVCISSILPDGEGEPATANGSKASKAPGLPEPEPE 388
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 326 VVLAVALSRKGMYPVRYNSPTFEMVCISSILPDGEGEPATANGSKASKAPGLPEPEPE 383
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 389 DREGDDLTLSFLP 402
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 384 DRDGDILTLSFLP 397
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
SGS3_DROME STANDARD; PRT; 307 AA.
ID SGS3_DROME
AC P02840; Q9VTJ2;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Saliivary glue protein Sgs-3 precursor.
GN SGS3 OR CG11720.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephyroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RX MEDLINE=83294545; PubMed=6411930;
RA Garfinkel M.D., Pruitt R.E., Meyerowitz E.M.;
RT "DNA sequences, gene regulation and modular protein evolution in the
RT Drosophila 68C glue gene cluster.";
RL J. Mol. Biol. 168:765-789(1983).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dong L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Mei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jialali M., Kalush F., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiter K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
RN SEQUENCE OF 1-28 FROM N.A.
RP MEDLINE=88332966; PubMed=3138416;
RX Martin C.H., Mayeda C.A., Meyerowitz E.M.;
RT "Evolution and expression of the Sgs-3 glue gene of Drosophila.";
RL J. Mol. Biol. 201:273-287(1988).
[4]
RN DEVELOPMENTAL STAGE.
RP MEDLINE=94038699; PubMed=8223281;
RX Huef F., Ruiz C., Richards G.;
RT "Puffs and PCR: the in vivo dynamics of early gene expression during
RT ecdysone responses in Drosophila.";
RL Development 118:613-627(1993).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: SALIVARY GLAND SPECIFIC.
CC -1- DEVELOPMENTAL STAGE: IN THE SALIVARY GLANDS OF MID INSTAR LARVAE
CC LEVELS DRAMATICALLY INCREASE DURING PUFF STAGE 1 AT 98-106 HOURS
CC OF DEVELOPMENT. LEVELS REMAIN CONSTANT AND ABUNDANT IN LATE LARVAE
CC UNTIL PUFF STAGE 10, THEN DECREASE BY STAGE 11.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC
CC EMBL, X01918; CAA25994.1; -.
CC EMBL, AE003544; AAF50056.1; -.
CC EMBL, X78392; CAA55154.1; -.
CC PIR, A03329; GSPF3.
CC FLYBASE: FBgn003373; Sgs3.
CC Repeat; signal.
CC SIGNAL 1 23 POTENTIAL.
CC CHAIN 24 307 SALIVARY GLUE PROTEIN SGS-3.
FT

```


SEQUENCE 307 AA; 32196 MW; 45803DED16C418BC CRC64;
Query Match 14.2%; Score 288.5; DB 1; Length 307;
Best Local Similarity 30.8%; Pred. No. 8.4e-11;
Matches 69; Conservative 33; Mismatches 112; Indels 11; Gaps 2;
QY 55 PETEPPMLNSTDTPPLTGPGESESTVEPA-----RSTGLDAGAVTELTTEL 106
DB 37 PRITQPECTTTTITTTTTCAPQOSTQTOPPCTTSKPTTKPTTKQTQLP---CTPTTKA 93
QY 107 ANNUNSTDSAMEIOTOPAAEAOPTPLAAEAOPTRLAAEAOPTPLAAEAOPTP 166
DB 94 TTTKPTTKATTTKATTKATTKPTTKQTQLPCTPTTKQTQLPCTPTTKPTTK 153
QY 167 AATEAOTTOPGLEAOTPAEAOPTPAEAOPTPAEAOPTPAEAOPTTAAEAOPTAP 226
DB 154 TTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTK 213
QY 227 EATEAOTTOPPAEAOPTPAEAOPTPAEAOPTPAEAOPTPAEAOPTPAEAOPT 270
DB 214 TTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTK 257

RESULT 4

VGKX_HSVB STANDARD; PRT; 797 AA.
AC P28968;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Glycoprotein X precursor.
GN 71.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86664; AAB02506.1; -
DR PIR; H36802; VGBEX1.
KM Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 797 GLYCOPROTEIN X.
FT DOMAIN 23 465 SER/THR-RICH.
FT TRANSMEM 766 790 POTENTIAL.
FT CARBOHYD 590 590 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 797 AA; 80342 MW; 50C9BD9211F5B552 CRC64;

Query Match 12.3%; Score 250.5; DB 1; Length 797;
Best Local Similarity 28.0%; Pred. No. 4.3e-08;
Matches 104; Conservative 33; Mismatches 169; Indels 65; Gaps 13;

QY 55 PETEPPMLNSTDTPPLTGPGESESTVEPA-----RSTGLDAGAVTELTTELANNGL 112
DB 53 PTTSPP-----TTSSSPSTSTHTSSSPSTSTOSSSTATSSAPSTASS-TTSTPTSTST 106
QY 113 STDSAMEIOTOP-----AATEAOTTOPAAEAOPTPAEAOPTPAEAOPTPAEAOPT 155
DB 107 ETTTTPASTTPTTTTAAPTTAATTAAVTAATTAATTAATTAATTAATTAATTAATTA 166

QY 156 -----LAATEAOTTPPAEAOPTTOPGLEAOTPAEAOPTPAEAOPTPAEAOPTP 205
DB 167 TTTATTTVPTASTTTTDTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 226
QY 206 PAEAOPTTOPPAEAOPTPAEAOPTPAEAOPTPAEAOPTPAEAOPTPAEAOPTPAE 264
DB 227 SATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 286
QY 265 EPTTKGLPIPRSVSVTHKGIIPMAA---SNLSVNVPGAP--DHISVQCLAILILAL 319
DB 287 TSTS-----TSTSAATTSPTPTSAATSAESTTEAPSTPTTDTTTPSE-----AT 333
QY 320 VATTFVCTVLAIVLRKRGHVPVNVSPTEMCISILPDGSGPATANGSLKAKS 379
DB 334 TATTSSESTVSASTSA-----TTTAFTESSHSPDSSTGSGSTSTABSSSTFLT 383
QY 380 PGLTPEPRDR 390
DB 384 PS-TATPSTDQ 393

RESULT 5

MUC2_HUMAN STANDARD; PRT; 5179 AA.
ID MUC2_HUMAN
AC 002817; Q14878;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mucin 2 precursor (Intestinal mucin 2).
GN MUC2 OR SMUC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=94132002; PubMed=8300571;
RA Gum J.R., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;
RT "Molecular cloning of human intestinal mucin (MUC2) cDNA.
RT Identification of the amino terminus and overall sequence similarity
RT to prepro-von Willebrand factor.";
RL J. Biol. Chem. 269:2440-2446(1994).
RN [2]
RP SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=93016075; PubMed=1400449;
RA Gum J.R., Hicks J.W., Toribara N.W., Roche E.-M., Lagace R.E.,
RA Kim Y.S.;
RT "The human MUC2 intestinal mucin has cysteine-rich subdomains located
RT both upstream and downstream of its central repetitive region.";
RL J. Biol. Chem. 267:21375-21383(1992).
RN [3]
RP SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
RX MEDLINE=9158717; PubMed=1885763;
RA Toribara N.W., Gum J.R., Culhane P.J., Lagace R.E., Hicks J.W.,
RA Petersen G.M., Kim Y.S.;
RT "MUC-2 human small intestinal mucin gene structure. Repeated arrays
RT and polymorphism.";
RL J. Clin. Invest. 88:1005-1013(1991).
CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
CC OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
CC AGENTS AT MUOCOSAL SURFACES.
CC -1- SUBUNIT: MULTIMERIC.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
CC BRONCHUS, CERVIX AND GALL BLADDER.
CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -1- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
CC VARIES AMONG DIFFERENT ALLELES.
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT

Query Match	Best Local Similarity	9.7%; Score 196.5; DB 1; Length 5179;
Matches 89; Conservative 24; Mismatches 169; Indels 81; Gaps 13;		
SO SEQUENCE	5179 AA; 540295 MM; 85CD751FB9A5663 CRC64;	
FT CARBOHYD	1139 1139	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	1154 1154	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	1215 1215	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	1230 1230	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	1246 1246	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	1787 1787	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	1820 1820	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4339 4339	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4351 4351	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4362 4362	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4373 4373	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4422 4422	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4438 4438	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4502 4502	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4616 4616	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4627 4627	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4752 4752	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4787 4787	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4881 4881	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4888 4888	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4955 4955	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	4970 4970	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	5019 5019	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	5038 5038	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CARBOHYD	5069 5069	N-LINKED (GLCNAC. .) (POTENTIAL.)
FT CONFLICT	1351 1351	H -> L (IN REF. 3).
FT CONFLICT	1412 1412	T -> S (IN REF. 3).
FT CONFLICT	1449 1449	L -> P (IN REF. 3).
FT CONFLICT	1504 1504	M -> T (IN REF. 3).
FT CONFLICT	4192 4192	G -> S (IN REF. 2).
QY 55 PETPEPMLKRNSTDTTPELGTGCTPESTVEPAARSTGLDAGAVLTETTLAMGNLST	114	
DB 3966 PTTTPPTTTTTVTTPPTPTGTGQTPTTTTTVTTPPTPTGTGQTPTTTTTVT-	4024	
OY 115 DSAMEIQTOPTATEA-QTPLPLATEAQTRLPLATEAQ--TTPLATEAQTTPATE	170	
DB 4025 -----PTPTPTGTGQTPTTTTTVTTPPTPTGTGQTPTTTTTVTTPPTPTG	4077	
OY 171 AQTTPGTLEAQT-----APAAEAQTTPAAMEAQTTP-PAAMEAQTTPAMEAQT	224	
DB 4078 TQTPTTTPTTTTTVTTPPTPTGTGQTPTTTTTVTTPPTPTGTGQTPTTTTT	4137	
OY 225 ----APBAEAQTOTPTATEAQTTPPLAMEALSTESBAEALSMETTKGLPIPSVS	279	
DB 4138 VTPTPTPTGTGQTPTTTTTVTTP-----TPPTPTGTGQTPTTTTTVTTPPTPT	4190	
OY 280 SVTHKGI PMAS-----NLSTNYVVGADHISVQCGLALILALVATIFVCIVLAVR	334	
DB 4191 TGTGTGPTPTTSTVPIALBLTISNP--PPSSSTPO-----	4222	
OY 335 LSRKGNHYPVRYNSP-TEMVCISILPLDGGEGSP-----ATANGLSKAKSPGLT	383	
DB 4223 TSRS-----TSSPLTSTLTLSTLPLALEMTSTAPSTPTATPTTSGHTLSRPSTT	4275	
OY 384 PEP 386		
DB 4276 TSP 4278		
RESULT 6		
DANA_YEAST STANDARD; PRT; 1161 AA.		
AC P4179;		
DT 01-FEB-1996 (Rel. 33, Created)		
DT 01-FEB-1996 (Rel. 33, Last sequence update)		
DT 15-JUN-2002 (Rel. 41, Last annotation update)		
DE Cell wall protein DANA precursor.		

GN DAN4 OR YJRI51C OR J2223.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA Scarsen T.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REGULATION.
 RX MEDLINE=21113168; PubMed=11160904;
 RA Cohen B.D., Settil O., Abramova N.E., Davies K.J., Lowry C.V.;
 RT "Induction and repression of DAN1 and the family of anaerobic
 RT manoprotein genes in Saccharomyces cerevisiae occurs through a
 RT complex array of regulatory sites."
 RL Nucleic Acids Res. 29:799-808(2001).
 CC -1- FUNCTION: COMPONENT OF THE CELL WALL (By similarity).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
 CC (potential).
 CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE SRP1 / TIPI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z49651; CAAB9684.1; -
 DR SGD; S0003912; DAN4.
 DR InterPro; IPR000992; SRP1_TIP1.
 DR Pfam; PF00660; SRP1_TIP1.1.
 DR PROSITE; PS00724; SRP1_TIP1; 1.
 KW Cell wall; Glycoprotein; Membrane; GPI-anchor; signal.
 DR FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 1146 CELL WALL PROTEIN DAN4.
 FT PROPEP 1147 1161 REMOVED IN MATURE FORM (POTENTIAL).
 FT LIPID 1146 1146 GPI-ANCHOR (POTENTIAL).
 SQ SEQUENCE 1161 AA; 118358 MW; 7954C15D65F0CA58 CRC64;
 Query Match 9.6%; Score 194; DB 1; Length 1161;
 Best Local Similarity 26.6%; Pred. No. 0.00015;
 Matches 73; Conservative 30; Mismatches 125; Indels 46; Gaps 11;
 QY 56 ETEPEPEMKNSTP---TPPLTG-----PGTP-ESTVEPAARRSTGLDAGAVT 100
 DB 62 ETPSEIATAAVFYGPFTRRLTGISGDEVTRMITGVWYSTRLKPAI--SSALSKDGIYT 119
 QY 101 ELTTELAMGNLSTDSAMBIQTTPAPAEQAOTPLAATAQAOTRLATQAOTPLAATE 160
 DB 120 AITP---STSTTKSTSTSTPTTITSTSTSTPTTSTSTPTTSTSTPTTST- 175
 QY 161 AQTTPPAATEA-----QTTPGLEAQTAPAPAMEAQTAPAPAMEAQTTPPAMEAQTQ 215
 DB 176 TSTPTTSTSTPTTSTPTTSTPTTSTSTPTTSTSTPTTSTPTTSTSTPTTST- 228
 QY 216 TTAEEAQTAPAEATEAQTTPAETEAQTPLAA---MEALSTEP-----SATEALS 263
 DB 229 TTPPTSTSTSTQSTKSTPTTSTSTPTTSTSTPTTSTSTPTTSTSTPTTSTSTSTIS 288
 QY 264 MEPTTKRGLFIPSVSVTHKGIIPMAASNSVNY 297
 DB 289 TAPTST-TSTPTSTSSASASVISTATSTT 321

RESULT 7
 YS89 CAEEL
 ID YS89 CAEEL STANDARD; PRT; 3178 AA.
 AC Q09624; Q09625; Q09694;
 DT 01-NOV-1995 (Rel. 32, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK945.9 in chromosome II.
 GN ZK945.9/ZK945.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RA SEQUENCE FROM N.A.
 RA STRAIN=Briscot N2;
 RA Wilkinson-Sprat U.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z48544; CAB70192.1; -
 DR EMBL; Z48582; CAB70192.1; JOINED.
 DR EMBL; Z48544; CAB70201.1; -
 DR WormRep; ZK945.9; CE25697.
 DR InterPro; IPR002111; Cat_channel_Tyrl.
 DR InterPro; IPR01024; Lipoxigenase_LH2.
 DR InterPro; IPR000636; M+channel_nlg.
 DR Pfam; PF00520; Ion_transf; 1.
 DR Pfam; PF01477; PLAT; 1.
 DR Pfam; PRO1825; GPS; 1.
 DR SMART; SM00303; GRS; 1.
 DR SMART; SM00308; LH2; 1.
 DR Hypothetical protein; Transmembrane.
 FT DOMAIN 266 1196 SER/THR-RICH.
 FT DOMAIN 1105 1241 GY/SER-RICH.
 FT DOMAIN 2071 2120 GPS.
 FT TRANSMEM 13 30 POTENTIAL.
 FT TRANSMEM 51 73 POTENTIAL.
 FT TRANSMEM 2139 2161 POTENTIAL.
 FT TRANSMEM 2348 2367 POTENTIAL.
 FT TRANSMEM 2390 2412 POTENTIAL.
 FT TRANSMEM 2451 2468 POTENTIAL.
 FT TRANSMEM 2483 2505 POTENTIAL.
 FT TRANSMEM 2567 2589 POTENTIAL.
 FT TRANSMEM 2836 2858 POTENTIAL.
 FT TRANSMEM 2939 2961 POTENTIAL.
 FT TRANSMEM 2976 2998 POTENTIAL.
 FT TRANSMEM 3038 3060 POTENTIAL.
 SQ SEQUENCE 3178 AA; 344726 MW; F8239436D03666CD CRC64;
 Query Match 9.5%; Score 193.5; DB 1; Length 3178;
 Best Local Similarity 22.6%; Pred. No. 0.00048;
 Matches 87; Conservative 57; Mismatches 170; Indels 71; Gaps 12;
 QY 65 NSTDTT---PLTGPSTVEPAARRSTGLDAGAVTELTELAMGN-----LSTD 115
 DB 336 SSTSTQSSSTSTSSPSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 388
 QY 116 SAAMEIQTQ--PAATEAQTPLAATAEAQT--RLTAEQAOTPLAATAEAQTTPAATEA 171
 DB 389 TTTSTTTTMTLSTTTEPSTSTTTTEVTSTSTSTSTSTSTSTSTSTSTSTSTST 448
 QY 172 QTTPGLEAQTAPAPAMEAQT--TAPAMEAQTTPPAMEAQTQTTPAMEAQTAPAEAT 229

```

Db 449 VTTSPSTSPVTSVTSSTSSSTVTTPTSTSTSTSPSS--TVTSTTAPSTSTTGPSS 506
Qy 230 BAQTOPTATBAQTPATLAMEALSTEPSATLMEPTTKRGFLFIPF-SVSSVTHKIGIM 288
Db 507 SSTPSTASSTSSSTSSSTQSTSTQSTSTTKSETTSDGNPFYVEKATTTFYDS 566
Qy 289 AASNLSTVNPVGPAPHISVKQCLLALLLALVATTFVCTVVLAVLRSKGMYP----- 343
Db 567 TSVNLTNLNGLGIGIGYQTSIECTSPSSNVSTTKGAC-FTKSVEMPLGGTYFPASTV 625
Qy 344 -----VKNYSPTEWVCISLLPDGSGSAPANGLSKA 377
Db 626 GGNVTFPATWTTDDKRVYTYANVYIOEYSSFTTIESSS-----TSAAVSSSTSTP 677
Qy 378 KSPGLT-----PEPREDEGDLT 396
Db 678 STPSSSTLSTSTVTEPSTSTSSDSTT 702

```

RESULT 8

AMVH YEAST STANDARD; PRT; 1367 AA.

```

ID AMVH YEAST STANDARD; PRT; 1367 AA.
AC P08640; P08068;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucanase 51/52 precursor (EC 3.2.1.3) (Glucan 1,4-alpha-
GN glucosidase) (1,4-alpha-D-glucan glucanhydrolase).
STAL OR STAZ OR MAL5 OR YIR019C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxId=4932;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Hornell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE OF 1-242 AND 762-1331 FROM N.A.
RC MEDLINE=87194600; PubMed=3106330;
RA Yamashita I., Nakamura M., Fukui S.;
RA "Gene fusion is a possible mechanism underlying the evolution of
RT STAL."
RT J. Bacteriol. 169:2142-2149(1987).

```

RN [3]

RP SEQUENCE OF 1-31 FROM N.A.

```

RC STRAIN=SPX101-1C;
RA MEDLINE=89031230; PubMed=3141213;
RA Paro J.M., Ianez E., Zalacain M., Claros M.G., Jimenez A.;
RT "Similar short elements in the 5' regions of the STR2 and SGA genes
RT from Saccharomyces cerevisiae."
RL FEBS Lett. 239:1175-1184(1988).

```

CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal 1,4-linked alpha-D-glucose residues successively from non-reducing ends of the chains with release of beta-D-glucose.

CC -1- SIMILARITY: TO S.POMBE SPBC215.13.

CC -1- SIMILARITY: SOME, TO S.POMBE SPCC285.13C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DR EMBL; Z38061; CA86176.1; -
DR EMBL; M16164; AAA5014.1; -
DR EMBL; M16165; AAA5015.1; -
DR EMBL; X13857; CAA32069.1; -
DR PIR; B26877; B26877.
DR PIR; A26877; A26877.
DR PIR; S48478; S48478.
DR SGD; S0001458; M0C1.
KW Hydrolase; Glycosidase; Polysaccharide degradation; Glycoprotein;
KW Signal; Multigene family.
FT SIGNAL 1
FT CHAIN 21
FT DOMAIN 210 1367
FT CARBOHYD 817 817
FT CARBOHYD 874 874
SQ SEQUENCE 1367 AA; 136110 MW; 91C00E2B9D51AA9D CRC64;

```

Query Match 9.2%; Score 187; DB 1; Length 1367; Best Local Similarity 24.0%; Pred. No. 0.00048; Matches 55; Conservative 36; Mismatches 110; Indels 28; Gaps 4;

```

Qy 76 GPRESTVTPARARSTGLDAGAVTELTTELAMGNLSSTDSAMEIQTTPATEAQTTP 135
Db 207 GTKSSTTSSSTSSST-----TTSSTSSSTTSSSTSSSTT 244
Qy 136 LAATEAQTTLTATEAQTTPPLATEAQTTPATEAQTTPGLAQTTPAPAMEAQTTPA 195
Db 245 SSTSSSTSSSTTATATTTTSCTEKRPPTTCTEKRPPTHDTT-PTCKKTTTS 303
Qy 196 PAMEAQTTPPAMEAQTTPQTAMEAQTTPATEAQTTPA--TEAQTPLAMEALS 253
Db 304 KTCTKTTTTPVTPPSSSTSSSAPVTPPSSSTSSSAPVTSSTSSSAPVTPPSSST 363
Qy 254 TEPATLMEPTTKRGFLFIPSVSVTHKGIPIMAASLSTNYVVGAP 302
Db 364 TESSAPVTS--STSSSAPVTSSTSSSAPVTPPSSSTSSSAP 409

```

RESULT 9

M0C1_XENLA STANDARD; PRT; 662 AA.

```

ID M0C1_XENLA STANDARD; PRT; 662 AA.
AC Q05049;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integumentary mucin C.1 (FIM-C.1) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxId=8355;

```

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 7).

```

RC TISSUE=Skin;
RA MEDLINE=93077556; PubMed=1447205;
RA Hauser F., Hoffmann W.;
RT C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
RT polymorphism."
RL J. Biol. Chem. 267:24620-24624(1992).

```

CC -1- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- ALTERNATIVE PRODUCTS: At least 7 isoforms; 1 (shown here), 2, 3, 4, 5, 6 and 7; may be produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: SKIN.

CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED.

CC -1- SIMILARITY: CONTAINS 6 P-TYPE (TREPOIL) DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X04798; CA28490.1; -
DR EMBL; Z86099; CAB06711.1; -
DR PIR; C43674; C43674.
DR InterPro; IPR003363; Herpes_gg.
DR Pfam; PF02400; Herpes_gg; 1.
KW Glycoprotein.
FT CARBOHYD 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 699 AA; 72243 MW; 612BA7B345E75540 CRC64;

Query Match
Best Local Similarity 8.7%; Score 176.5; DB 1; Length 699;
Matches 95; Conservative 49; Mismatches 167; Indels 105; Gaps 17;

QY 24 WA---DEAEKALGPLARDRQ-ATEYEVLDYDF-----LPETPEEMLRNST 67
DB 300 WAPGALDDGYPAPPPRRFRRLRTDPEGVDVAPRTGRRLMTEDTSDSPTSAP 359
QY 68 DTPLPLGPGTPESTYEPARRSTGLDAGAVT---ELTTELANMGNLSDSAMEIQT 124
DB 360 EKPLPLVSATAMPSVDSAEPT---APATTPPEMATQAATVATPEETIV---AS 411
QY 125 QPAATEAQTPLAATAQT-----TRLTATEAQTPLAATAQTTPAATAQTPTGL 179
DB 412 PPAATAVESSEPLAAAAATPGAGHTNTSSAAKTP-PTTPAATTPPTTSHATPTPT 470
QY 180 EAQTT-----APAAEAOQTAPAAEAOQTTPAAEAOQTTPAAEAOQTTPAAEAO 232
DB 471 GPOTTPPGPATPGVGSAPATPDSPLTA--SPAPATPGPSANVVAATATPTGRGTA 528
QY 233 TTPQT-----ATEAQTPLAMEAL 252
DB 529 RPPPTDPKTHPHGPADAPGSSPAPPPPEHNGEPEEGAGDGPEDDDSATLARPTN 568
QY 253 STEPSATEALSMETTKRGFLIFPSVSVTHKGIIPMAASNLVNVFVG-APDHISYKQCL 311
DB 589 PKNPPPARPGPIRPTLPGLGLGLAPNTPR---PRAQAP-AKDMPSGPIPHIPLFWL 643
QY 312 LA-----LILALVATTFVCTVLAARL-----SRKGHMYPVNNYSPTENVC 355
DB 644 TASPALDILFIISTTHTAFCVLAALQLMRGRAGRRRYAPSVRY-----VCL 694

RESULT 13
CH12_COCIM STANDARD; PRT; 860 AA.
ID CH12_COCIM
AC P54197;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Endochitinase 2 precursor (EC 3.2.1.14).
GN CTS2.
OS Cocciidioides immitis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Onygenales; mitosporic Onygenales; Cocciidioides.
OX NCBI_TaxID=5501;
RN 11
RP SEQUENCE FROM N.A.
BC STRAIN=C735;
RX MEDLINE=96144270; PubMed=8566773;
RA Bishop E.J., Kirkland T.N., Cole G.T.;
RT "Isolation and characterization of two chitinase-encoding genes
CT (ctsl, cts2) from the fungus Cocciidioides immitis.";

RL Gene 167:173-177(1995).
CC -1- FUNCTION: MAY BE ASSOCIATED WITH ENDOSPORULATION.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
CC acetyl-D-glucosamine polymers of chitin.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L41662; AAA92642.1; -
DR HSSP; P23472; 2HW.
DR InterPro; IPR001579; Chitinase_18/2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR Prosite; PS01095; CHITINASE_18; 1.
KW Hydrolyase; Glycosidase; Chitin degradation; Chitin-binding; Signal;
KW Glycoprotein.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 860 ENDOCHITINASE 2.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 860 AA; 91395 MW; 5E34B54FA63F3C CRC64;

Query Match
Best Local Similarity 8.5%; Score 173.5; DB 1; Length 860;
Matches 82; Conservative 53; Mismatches 141; Indels 93; Gaps 14;

QY 12 LCPGNSIQIMPTMADEAEKALGPLARDRQATEYEVLDYDFP-----ETEP 59
DB 288 LPPGENTKIVSYMAKYPSTFGMV---WEATBAENNKGLPYADIMKEVLACDPP 344
QY 60 PEMLRNST-----DTPLTGPPTES-TTVEPARRSTGLDAGAVTEL 102
DB 345 PNSTVSTTSASTSQTSSQSTTMTKTLASTPSSPSSTVMTQSTTSGSTIEIV 404
QY 103 TTELAN--MGNLSTSSAMEIQTTPAATAQTTP-----LAATEAQTTL 146
DB 405 TTRSQEPSTTISTSA-----STEPVTRSQEPSTTISTRSASTETVTRSQEPSTT 459
QY 147 -----TATEAQTTP-----LAATEAQTTPAATAQTTP 176
DB 460 ISTWSASTSTSSQDSPTTISTSAPIGYTTTSQDLPTTISTRSSTETTRATTKS 519
QY 177 TGLEAQTAPAAEAOQTAPAAEAOQTTPAAEAOQTTPAAEAOQTTPAAEAOQTTP 236
DB 520 QGSPSTILSTRSSSAETVSTRSQHSSSTTISTKSAPTETGTSTSEHSTSPVSTRASST 579
QY 237 TATEAQTPLAMEALSTEPSATEALSMETTKRGFLIFPSVSVTHKGIIPMAASNLSTN 296
DB 580 VITRQNSDSQSM-TVSTRSPSTESI-----TTRSQGSPSEFTST--KSVPV--DTISTE 629
QY 297 YPVGAPDHI 305
DB 630 LPSQTPPTI 638

RESULT 14
ZAN_MOUSE
ID ZAN_MOUSE STANDARD; PRT; 5376 AA.
AC O88799; O08647;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zonadhesin precursor.
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98123114; PubMed=9452463;
RA Gao Z., Garbers D.L.;
RT "Species diversity in the structure of zonadhesin, a sperm-specific
RT membrane protein containing multiple cell adhesion molecule-like
RT domains";
RT J. Biol. Chem. 273:3415-3421 (1998).
RN [2]
RP SEQUENCE OF 4864-5376 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=97271566; PubMed=9126492;
RA Gao Z., Hatumi T., Garbers D.L.;
RT "Chromosome localization of the mouse zonadhesin gene and the human
RT zonadhesin gene (ZAN).";
RT Genomics 41:119-122 (1997).
CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC SIGNALING.
CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC APICAL REGION OF THE SPERM HEAD.
CC -1- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
CC -1- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
CC ZONA PELLUCIDA.
CC -1- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
CC SPERMATOZOON OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
CC -1- DOMAIN: THE VWFD DOMAIN 2 MAY MEDIATE COVALENT
CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
CC -1- SIMILARITY: CONTAINS 3 MAM DOMAINS.
CC -1- SIMILARITY: CONTAINS 25 VWFD DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; U97068; AAC26680.1; -
CC MGD; MGI:106566; Zm.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR0003645; FOLN.
DR InterPro; IPR000998; MAM domain.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR003328; TIL_Cysrich.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00094; vwd; 4.
DR Pfam; PF00629; MAM; 3.
DR Pfam; PF01826; TIL; 25.
DR Pfam; PF02345; TILa; 25.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00274; FOLN; 11.
DR SMART; SM00137; MAM; 2.
DR SMART; SM00214; VMC; 17.
DR SMART; SM00216; VMD; 4.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 18.
DR PROSITE; PS00740; MAM_1; FALSE_NEG.
DR PROSITE; PS00060; MAM_2; 3.
KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
KW Repeat.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 5376 ZONADHESIN.
FT DOMAIN 18 5310 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 5311 5337 POTENTIAL.

FT	DOMAIN	5338	5376	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	45	210	MAM 1.
FT	DOMAIN	215	374	MAM 2.
FT	DOMAIN	377	542	MAM 3.
FT	DOMAIN	547	1170	60 X HEPTAPEPTIDE REPEATS (APPROXIMATE (MUCIN-LIKE DOMAIN).
FT	DOMAIN	1171	1280	VWFD 1 (PARTIAL).
FT	DOMAIN	1281	1669	VWFD 2.
FT	DOMAIN	1670	2056	VWFD 3.
FT	DOMAIN	2057	2459	VWFD 4.
FT	DOMAIN	2460	2579	VWFD 5 (PARTIAL).
FT	DOMAIN	2580	2639	VWFD 6 (PARTIAL).
FT	DOMAIN	2700	2819	VWFD 7 (PARTIAL).
FT	DOMAIN	2820	2939	VWFD 8 (PARTIAL).
FT	DOMAIN	2940	3059	VWFD 9 (PARTIAL).
FT	DOMAIN	3060	3179	VWFD 10 (PARTIAL).
FT	DOMAIN	3180	3299	VWFD 11 (PARTIAL).
FT	DOMAIN	3300	3416	VWFD 12 (PARTIAL).
FT	DOMAIN	3417	3536	VWFD 13 (PARTIAL).
FT	DOMAIN	3537	3656	VWFD 14 (PARTIAL).
FT	DOMAIN	3657	3776	VWFD 15 (PARTIAL).
FT	DOMAIN	3777	3892	VWFD 16 (PARTIAL).
FT	DOMAIN	3893	4928	VWFD 17 (PARTIAL).
FT	DOMAIN	4029	4148	VWFD 18 (PARTIAL).
FT	DOMAIN	4149	4263	VWFD 19 (PARTIAL).
FT	DOMAIN	4264	4283	VWFD 20 (PARTIAL).
FT	DOMAIN	4384	4503	VWFD 21 (PARTIAL).
FT	DOMAIN	4504	4623	VWFD 22 (PARTIAL).
FT	DOMAIN	4624	4743	VWFD 23 (PARTIAL).
FT	DOMAIN	4744	4863	VWFD 24 (PARTIAL).
FT	DOMAIN	4864	5261	VWFD 25.
FT	DOMAIN	5259	5295	EGF-LIKE.
FT	DISULFID	5263	5274	BY SIMILARITY.
FT	DISULFID	5268	5283	BY SIMILARITY.
FT	DISULFID	5285	5294	BY SIMILARITY.
FT	CARBOHYD	339	339	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	499	499	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1216	1216	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1239	1239	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1314	1314	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1814	1814	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1908	1908	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	1933	1933	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2028	2028	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2111	2111	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2142	2142	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2332	2332	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2533	2533	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2575	2575	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2692	2692	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	2812	2812	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3052	3052	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3065	3065	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3144	3144	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3172	3172	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3288	3288	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3292	3292	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	3782	3782	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4005	4005	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4136	4136	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4243	4243	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4254	4254	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4335	4335	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4376	4376	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	4586	4586	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	5136	5136	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	5252	5252	N-LINKED (GLCNAc. . .) (POTENTIAL).
SO	SEQUENCE	5376 AA;	579908 MW;	0E44DB77DF2A2620 CRC64;

Query Match 8.4%; DB 1; Length 5376;
Best Local Similarity 28.6%; Pred. No. 0.018;
Matches 66; Conservative 24; Mismatches 107; Indels 34; Gaps 9;


```

QY 57 TEBPEMLRNSTDTPLTGPSTPESTVEPARRSTGLDAGAVTELTTELAMNGNISTDS 116
Db 948 TLFTLEVL-----TVPLEVTTFPGTEVTVPTEVTSTEMGVHTEVTVEPSTISPT-- 1000
QY 117 AMMEIOTTOPA--TEAQTPLATEAQTTRLTAT---EAQTPLAA-----TEA 161
Db 1001 ---EVAIVLPASIPPEETITP---TEVTTPPEETIPAVETVPPASIPPEETASITEV 1054
QY 162 OTTPPATEAQTQPTGLEAQTAPAMEAQTAPAM--EAQTTPPAMEAQTOTTAM 219
Db 1055 TTTPPEETITPTEVTVPPEETITPT--EVTVPASIPPEETVPEETITASEETVS 1112
QY 220 EAQTT--APPEATEAQTQPTATEAQTPLAMEALSTEPSATEALSMEPTT 268
Db 1113 TOSTTLLTEQSATVOTSIACRPPCPSPPLMIPGLSKPVGWMSFIAPFTT 1163

RESULT 15
HKRI YEAST STANDARD; PRT; 1802 AA.
ID HKRI YEAST
AC P41809;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hansenula WRAKII killer toxin-resistant protein 1 precursor.
GN HKRI OR YDR420W.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxId=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YNN 295;
RX MEDLINE=94156857; Pubmed=8113191;
RA Kasahara S., Yamada H., Mio T., Shiratori Y., Miyamoto C.,
RA Yabe T., Nakajima T., Ichishima E., Furuchi Y.,
RA "Cloning of the Saccharomyces cerevisiae gene whose overexpression
RT overcomes the effects of Hm-1 killer toxin, which inhibits
RT beta-glucan synthesis."
RL J. Bacteriol. 176:1488-1499 (1994).
CC -1- FUNCTION: COULD REGULATE BETA-GLUCAN SYNTHESIS. OVEREXPRESSION
CC PROVIDES RESISTANCE TO HM-1 KILLER TOXIN.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -1- PTM: COULD BE O-GLYCOSYLATED IN SERINE/THREONINE RICH DOMAIN.
CC -1- SIMILARITY: SOME, TO YEAST MSB2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S69101; AAB30051.1; -.
DR SGD; S0002828; HKRI.
DR KW Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 1802 HANSENULA WRAKII KILLER TOXIN-RESISTANT
FT TRANSMEM 1486 1506 PROTEIN 1.
FT DOMAIN 23 1478 SER/THR-RICH.
FT 453 788 12 X 28 AA TANDEM REPEATS OF S-(AV)-[P]-
FT V-A-V-S-S-T-Y-T-S-S-P-S-A-P-A-A-I-S-S-T-
FT Y-T-S-S-P-
FT 453 480 1 (APPROXIMATE).
FT REPEAT 481 508 2.
FT REPEAT 509 536 3.
FT REPEAT 537 564 4.
FT REPEAT 565 592 5.
FT REPEAT 593 620 6.
FT REPEAT 621 648 7.
FT REPEAT 649 676 8.

```

```

FT REPEAT 677 704 9.
FT REPEAT 705 732 10.
FT REPEAT 733 760 11.
FT REPEAT 761 788 12.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1252 1252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1293 1293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1342 1342 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1400 1400 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1802 AA; 188890 MW; E344CA6469785a24 CRC64;

Query Match 8.3%; Score 169.5; DB 1; Length 1802;
Best Local Similarity 23.7%; Pred. No. 0.0071;
Matches 86; Conservative 63; Mismatches 145; Indels 69; Gaps 15;

QY 66 STDTPPLTGPSTPESTVEPARRSTGLDAGAVTELTTELAMNGNISTDSAMEIQTQ 125
Db 478 SSPSVPAVAVSSTYTSSPSAPAAISSTYTSSPSAPAVAVSSTYT---SSPSAPAAISSTY 532
QY 126 PATEAQTPLATEAQTTRLTAT---TEAQTPLATEAQTTPPATEAQTQPT 177
Db 533 ---TSPPSAPAVAVSSTYTSSPSAPAAISSTYTSSPSAPAVAVSSTYTSSPSAPAAISSTYT 589
QY 178 GLEAQTAPAMEAQT---APAMEAQT---PPAMEAQTOTTAMEA-----QTT 224
Db 590 ---SSPSAPAVAVSSTYTSSPSAPAAISSTYTSSPSAPAVAVSSTYTSSPSAPAAISSTYTSS 646
QY 225 APE---ATEAQTQPTA-----TEAQTPLAMEALSTEPSATEALSMEPTTKGL 272
Db 647 SSPSVPAVAVSSTYTSSPSAPAAISSTYTSSPSVPAVAVSSTYTSSPSAPAAISSTYTSSPS 705
QY 273 FIPFSVSVTHKGIIPMAASNLGVNY-----PV-----GAPDHISVQCLLAILI 316
Db 706 -APAVVSS--TYTSSPSAPAAISSTYTSSPSAPAVAVSSTYTSSPSAPAAISSTYTSSPSAP 763
QY 317 LALVATIFVCTVAVLAVLSRKGHMYPVANISPTTEWVCISLLPDGGEPSATANGGLSK 376
Db 764 VAVSSTYTSSPSALVLSSTSTSSPYDI-VVSPSTFAAIS-----GYTPSPSASVAMSS 817
QY 377 AKS 379
Db 818 TSS 820

```

Search completed: July 14, 2003, 06:15:13
 Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using SW model

Run on: July 14, 2003, 06:10:58 ; Search time 80 Seconds

(without alignments)
1035.387 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030
Sequence: 1 MPELQILLILGPGNSIQ.....TPPREDRGDDTLHSLFP 402

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	824.5	40.6	397	11	Q99L34
2	283	13.9	867	12	Q39782
3	271.5	13.4	866	12	Q39781
4	269	13.3	372	12	Q8VOM3
5	263	13.0	357	12	Q8VOM2
6	261.5	12.9	342	12	Q8VOM8
7	261	12.9	826	12	Q8VOM5
8	259	12.8	337	12	Q8VOM1
9	256.5	12.6	750	12	Q39307
10	256	12.5	356	12	Q8VOM7
11	253.5	12.3	1079	5	Q9N4S7
12	249.5	12.0	293	12	Q8VOM4
13	244	12.0	316	12	Q8VOM4
14	243.5	12.0	389	12	Q8VOM0
15	241	11.9	374	12	Q8VOM6
16	239.5	11.8			

17	232.5	11.5	703	12	Q8VOM3	Q8VOM3 equine herp
18	232.5	11.5	779	12	Q8VOM2	Q8VOM2 equine herp
19	227.5	11.2	273	12	Q8VOM0	Q8VOM0 equine herp
20	226.5	11.2	801	5	Q23635	Q23635 caenorhabdi
21	225.5	11.1	1195	3	Q96W08	Q96W08 schizosacch
22	222	10.9	804	12	Q8VOM1	Q8VOM1 equine herp
23	218.5	10.8	825	12	Q8VOM0	Q8VOM0 equine herp
24	218	10.7	291	12	Q8VOM8	Q8VOM8 equine herp
25	218	10.7	726	12	Q8VOM4	Q8VOM4 equine herp
26	211.5	10.4	240	12	Q8VOM5	Q8VOM5 equine herp
27	208.5	10.3	769	5	Q17921	Q17921 caenorhabdi
28	208	10.2	218	12	Q8VOM3	Q8VOM3 equine herp
29	207.5	10.2	260	12	Q8VOM6	Q8VOM6 equine herp
30	206	10.1	704	5	Q9GYA2	Q9GYA2 leishmania
31	205.5	10.1	1151	13	Q57580	Q57580 gallus gal1
32	204.5	10.1	770	5	Q20908	Q20908 caenorhabdi
33	199.5	9.8	668	5	Q9GY11	Q9GY11 leishmania
34	199	9.8	258	12	Q8VOM9	Q8VOM9 equine herp
35	198	9.8	3191	5	Q01335	Q01335 caenorhabdi
36	195	9.6	985	12	Q67643	Q67643 gallid herp
37	194.5	9.6	217	12	Q8VOM2	Q8VOM2 equine herp
38	194.5	9.6	245	12	Q8VOM7	Q8VOM7 equine herp
39	194.5	9.6	507	13	Q13028	Q13028 boreogadus
40	193.5	9.5	3178	5	Q969D4	Q969D4 caenorhabdi
41	193	9.5	658	3	Q59779	Q59779 schizosacch
42	193	9.5	1480	10	Q9LIE8	Q9LIE8 arabidopsis
43	191.5	9.4	5374	11	Q99ND0	Q99ND0 mus musculu
44	191	9.4	800	3	Q8RF64	Q8RF64 schizosacch
45	190.5	9.4	746	5	Q9V515	Q9V515 drosophila

ALIGNMENTS

RESULT 1
Q99L34 PRELIMINARY: PRT; 397 AA.
ID Q99L34;
AC Q99L34;
DT 01-JUN-2001 (TRENBLREL. 17, Created)
DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Similar to selectin, platelet (p-selectin) ligand.
GN SELPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003874; AA03874.1; --
DR MGD; MGI:106689; Selpl.
KW Lectin; Selectin.
SQ SEQUENCE 397 AA; 41844 MW; 552C7791622D23A3 CRC64;

Query Match 40.6%; Score 824.5; DB 11; Length 397;
Best Local Similarity 44.5%; Pred. No. 3.7e-42;
Matches 193; Conservative 53; Mismatches 119; Indels 69; Gaps 8;
QY 1 MPELQILLILGPGNSIQ...MDTWADBAKALGPIIARDRQAT-EYEYLDVFLPETERP 59
DB 1 MSPSPVLVLTITIGPNSIQ...DPWHEKKEAGPVHLERRRVVGDDEDDPDYTNIDP 60
QY 60 PEMLNSTIDTTP...TGPGTPESTVEPARSSTGIAGAVTELT 104
DB 61 PELIKAVNTVVAHDELPTVVMLERDSTSGTSERATEIKITPTIPAGTGTA----- 115
QY 105 ELAANGNISTSAAMEIQTTQPAATAEQTTPLAATAEQTRLTATAEQTTPLAATAEQTT 164
DB 116 ---VGMSTDSA-----TQMSLTSVETVQPASTEVEFS 145
QY 165 PPAATEAQTPTGIAQTAPAAEAQTAPAAEAQTTPPAEAQTTPPAEAQTTPPAEAQTT 224

```

Db      146 QPTPMADISKPAPMEAFETSQPAPMEAFETSQPAPMEAFETSQPAPMEAFETSQPAPMEAFETS 205
QY      225 APEATEAQTQPTATAEATQPTPL--AAMELSTEPSATELSNEPTTKGLFIPFSV-- 279
Db      206 KPAFPAETSKPAPAEATQTLQRIQAVKTLFTTSATEPSTPEPTMETASTESNESTI 265
QY      280 ----SVTH-----KGIPMAASNLNVYVGAAPDHISVKQCLLILILALVATIFVCT 328
Db      266 FLGSPVTHLPDGLKGLIVTPGSSAPRLPGSSDLIPKQCLLILILALVATIFVCT 325
QY      329 VVLAVALSRKGMYPVRNYSPTMVCISLLPDGSGPATANGLSKAKSPGLTEPRE 388
Db      326 VVLAVALSRKTHMYPVRNYSPTMVCISLLPDGSGPATANGLPKVD--LKTPESG 383
QY      389 DRGDDLTLSFLP 402
Db      384 DRGDDLTLSFLP 397

```

RESULT 2

```

039782 PRELIMINARY; PRT; 867 AA.
AC 039782;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Membrane glycoprotein.
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
OX NCBI_Taxid=10326;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BK343;
RA Kiriisawa R., Kobayashi T., Kawakami Y., Iwai H.;
RT "Nucleotide sequences of open reading frames 1, 24 and 71 of an
RT attenuated equine herpesvirus-1."
RL J. Equine Sci 7:79-87(1996).
DR EMBL; D88734; BAA20037.1; -.
SQ SEQUENCE 867 AA; 86630 MW; C8DDAD90D9B64407C CRC64;

```

Query Match 13.9%; Score 283; DB 12; Length 867;
 Best Local Similarity 31.2%; Pred. No. 2.5e-09;
 Matches 103; Conservative 23; Mismatches 174; Indels 30; Gaps 7;

```

QY      67 TDTPLTGPGTPESTVEPARRSSTGLDAGAVTELTTELANNGNLSTDSAMEIQTOP 126
Db      158 TTTTPTSTTTTATTTTPTTASTTTDTTATTAATTAATTAATTAATTAATTAATTA 217
QY      127 AATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLA 186
Db      218 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 277
QY      187 AAMEAQTPLAAMEAQTPLAAMEAQTPLAAMEAQTPLAAMEAQTPLAAMEAQTPL 246
Db      278 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 337
QY      247 AAMEALSTE-PSATEALSMEPTTKRGLFIPFSVSVTHKGIPIAA--SNLSVNYVGA 302
Db      338 TSTTGASTSTPSASTSATPTS-----TSTSAATTSPTPTPSAATSASTTEAPTSTP 392
QY      303 --DHISVKQCLLILILALVATIFVCTVVLAVRLSRKGMYPVRNYSPTMVCISLIP 360
Db      393 TDTTTPSE-----ATTATSPESSTVSASTSA-----TTTAFTTSHSTP 434
QY      361 DGEGSPATANGLSKAKSPGLTEPREDR 390
Db      435 DSSTGSTAEPSTFTLTPS-TATPSTDQ 463

```

RESULT 3

039781

```

ID 039781 PRELIMINARY; PRT; 866 AA.
AC 039781;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Membrane glycoprotein.
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
OX NCBI_Taxid=10326;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BH1;
RA Kiriisawa R., Kobayashi T., Kawakami Y., Iwai H.;
RT "Nucleotide sequences of open reading frames 1, 24 and 71 of an
RT attenuated equine herpesvirus-1."
RL J. Equine Sci 7:79-87(1996).
DR EMBL; D88733; BAA20037.1; -.
SQ SEQUENCE 866 AA; 86463 MW; 153BD142ABAAE88D CRC64;

```

Query Match 13.4%; Score 271.5; DB 12; Length 866;
 Best Local Similarity 30.9%; Pred. No. 1.2e-08;
 Matches 102; Conservative 23; Mismatches 170; Indels 35; Gaps 8;

```

QY      67 TDTPLTGPGTPESTVEPARRSSTGLDAGAVTELTTELANNGNLSTDSAMEIQTOP 126
Db      162 TTTTPTSTTTTATTTTPTTASTTTDTTATTAATTAATTAATTAATTAATTAATTA 216
QY      127 AATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPLAATEAQTPL 186
Db      217 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 276
QY      187 AAMEAQTPLAAMEAQTPLAAMEAQTPLAAMEAQTPLAAMEAQTPLAAMEAQTPL 246
Db      277 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 336
QY      247 AAMEALSTE-PSATEALSMEPTTKRGLFIPFSVSVTHKGIPIAA--SNLSVNYVGA 302
Db      337 TSTTGASTSTPSASTSATPTS-----TSTSAATTSPTPTPSAATSASTTEAPTSTP 391
QY      303 --DHISVKQCLLILILALVATIFVCTVVLAVRLSRKGMYPVRNYSPTMVCISLIP 360
Db      392 TDTTTPSE-----ATTATSPESSTVSASTSA-----TTTAFTTSHSTP 433
QY      361 DGEGSPATANGLSKAKSPGLTEPREDR 390
Db      434 DSSTGSTAEPSTFTLTPS-TATPSTDQ 462

```

RESULT 4

```

ID 08VOM3 PRELIMINARY; PRT; 372 AA.
AC 08VOM3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycoprotein gp2 (Fragment).
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
OX NCBI_Taxid=10326;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=974/91;
RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4)
RT and EHV1."
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY034637; AK61477.1; -.
DR InterPro; IPR000104; Anticfzeeze_1.
DR PRINTS; PR00308; ANTIFREZEI.
FT NON TER 372 372
SQ SEQUENCE 372 AA; 34391 MW; 5F2EBB27A5BFC41B CRC64;

```

Query Match 13.3%; Score 269; DB 12; Length 372;
 Best Local Similarity 37.5%; Pred. No. 7.2e-09;
 Matches 78; Conservative 9; Mismatches 115; Indels 6; Gaps 1;

QY 67 TDTPTGCTPESTVEPARRSGLDAGAVTELTTELAMNGNSTSAAEIOPTOP 126
 DB 158 TTTTPTSTTTTAAATTTTPTTASTTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 217
 QY 127 AATEAOTPLAATEAOTRLTATEAOTPLAATEAOTPPAATEAOTQPTGLEAOTAP 186
 DB 218 ATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTA 277
 QY 187 AAMEAQTAPAAEAOQTTPPAEAOQTTPPAEAOQTTPPAEAOQTTPPAEAOQTTP 246
 DB 278 ATTSATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 337
 QY 247 AAMEALST-----EPSATEALSMETT 268
 DB 338 PTSGSTSTTGASTSTPSASTATSATPTS 365

RESULT 5

Q8VOM2 PRELIMINARY; PRT; 357 AA.
 ID Q8VOM2; AC Q8VOM2; DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Glycoprotein gp2 (Fragment).
 OS Equine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae;
 OC NCBI_Taxid=10326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1026A/93;
 RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
 RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY034638; AK61478.1; -;
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PRO0308; ANTIFREEZE1.
 FT NON TER 357
 SQ SEQUENCE 357 AA; 33054 MW; 08310EF29F5B2641 CRC64;

Query Match 13.0%; Score 263; DB 12; Length 357;
 Best Local Similarity 33.3%; Pred. No. 1.6e-08;
 Matches 83; Conservative 11; Mismatches 139; Indels 16; Gaps 3;

QY 59 PREMLNSDTTPTLPGPPESTVEPARRSGLDAGAVTELTTELAMNGNSTDSAA 118
 DB 101 PTSTSTTTTPTTASTTTPTTTTAAATTTAATTTAATTTAATTTAATTTAATTTA 160
 QY 119 MEIOPTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTPPAATEAOT 173
 DB 161 TPTSTTTTATTTTPTTASTTTTDTTAAATTTAATTTAATTTAATTTAATTTAATTT 220
 QY 174 TOPTGLEAOTAPAAEAOQTTPPAEAOQTTPPAEAOQTTPPAEAOQTTPPAEAOQT 233
 DB 221 TAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 279
 QY 234 TQPTATEAOTPLAAMEALSTEPSATEALSMETTTRGIFIPSSVSHKGPMAANU 293
 DB 280 TAAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 329
 QY 294 SVNYPVGAP 302
 DB 330 TTGASTSTP 338

RESULT 6

Q8VOM8 PRELIMINARY; PRT; 342 AA.
 ID Q8VOM8; AC Q8VOM8; DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Glycoprotein gp2 (Fragment).
 OS Equine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae;
 OC NCBI_Taxid=10326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1077/94;
 RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
 RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY034642; AK61482.1; -;
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PRO0308; ANTIFREEZE1.
 FT NON TER 342
 SQ SEQUENCE 342 AA; 31718 MW; 3A0D3BE0DEF0D688 CRC64;

Query Match 12.9%; Score 261.5; DB 12; Length 342;
 Best Local Similarity 34.9%; Pred. No. 1.8e-08;
 Matches 82; Conservative 10; Mismatches 118; Indels 25; Gaps 3;

QY 59 PREMLNSDTTPTLPGPPESTVEP--AARSGLDAGAVTELTTELAMNGNST-- 114
 DB 101 PTSTSTTTTPTTASTTTPTTTTAAATTTAATTTAATTTAATTTAATTTAATTTA 160
 QY 115 -----DSAMEIOPTOPATEAOTPLAATEAOTRLTATEAOTPLAAT 159
 DB 161 TPTSTTTTATTTTPTTASTTTTDTTAAATTTAATTTAATTTAATTTAATTTAATTT 220
 QY 160 EAQTPPAATEAOTQPTGLEAOTAPAAEAOQTTPPAEAOQTTPPAEAOQTTPPAE 219
 DB 221 TAATTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 280
 QY 220 EAQTPPAATEAOTQPTPTEAOTPLAAMEALST-----EPSATEALSMETT 268
 DB 281 AATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 335

RESULT 7

Q8VOM5 PRELIMINARY; PRT; 826 AA.
 ID Q8VOM5; AC Q8VOM5; DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Glycoprotein gp2.
 OS Equine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirinae;
 OC NCBI_Taxid=10326;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=438/77;
 RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
 RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY034645; AK61485.1; -;
 DR InterPro; IPR000104; Antifreeze_1.
 DR PRINTS; PRO0308; ANTIFREEZE1.
 SQ SEQUENCE 826 AA; 82915 MW; F5B61157739CA74D CRC64;

Query Match 12.9%; Score 261; DB 12; Length 826;
 Best Local Similarity 29.4%; Pred. No. 5e-08;
 Matches 102; Conservative 25; Mismatches 180; Indels 40; Gaps 9;

```

QY 59 PPEMLNSTDTPLTGTGCTPESTTVEBPARRSTGLDAGAVTELTTELAMNGNSTDSAA 118
AC 101 PTTSTSTTTTTPASTTTPTTTTAAPTTAATTAATTAATTAATTAATTAATTAATTAAT 160
DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
QY 119 MEIOQT-----QPAATEAOTTPLAATEAOTRLTAATEAOTTPLAATEAOTTPAAT 169
DB 161 TPTSTTTTATTTTPTASTTTTDTTAAATTAATTAATTAATTAATTAATTAATTAAT 220
QY 170 EAOTTPGTGLEAOTTAAPAMEAOTTAAPAMEAOTTPPAAMEAOTTPPAAMEAOTTPAAT 229
DB 221 TAATTS-SATTAATTSATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 279
QY 230 EAOTTPAATEAOTTPLAAMEALSTE-PSATEALSMEPTTGGIFPVSVTHKGIM 288
DB 280 AATTTAATTTGSPSGSTSTGASTSPSASTATSAPTS-----TSTSAATTSPTPT 334
QY 289 AA---SLSLVNYPVGPAP--DHISVKQCLAILILALVATTFVCTVVLARLSRKHMP 343
DB 335 SAATSASTTEAPSTPTTDTTTPSE-----ATTATTSPESTTVSASTTSA----- 380
QY 344 VRNYSPTENVICISLDPGEGEPATANGGLSKAKSPGLTPEPREDR 390
DB 381 -----TTTATTTSTSTSPDSSGTSTABPSSTFLTTS-TATPTDQ 422

```

RESULT 8

```

Q8VOM1 PRELIMINARY; PRT; 337 AA.
ID Q8VOM1
AC Q8VOM1;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Glycoprotein gp2 (Fragment).
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae;
NCBI_TaxID=10326;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=1053A;
RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034639; AAK61479.1; -
DR InterPro; IPR000104; Antifreeze_1.
DR PRINTS; PR00308; ANTIFREEZE1.
FT NON TER 337
SQ SEQUENCE 337 AA; 31272 MW; 7771F755727A6E1C CRC64;

```

Query Match 12.8%; Score 259; DB 12; Length 337;
 Best Local Similarity 34.8%; Pred. No. 2.6e-08;
 Matches 80; Conservative 12; Mismatches 118; Indels 20; Gaps 3;

```

QY 59 PPEMLNSTDTPLTGTGCTPESTTVEBPARRSTGLDAGAVTELTTELAMNGNST-- 114
DB 101 PTTSTSTTTTTPASTTTPTTTTAAPTTAATTAATTAATTAATTAATTAATTAAT 160
QY 115 -----DSAMEIQTOTPAATEAOTTPLAATEAOTRLTAATEAOTTPAAT 159
DB 161 TPTSTTTTATTTTPTASTTTTDTTAAATTAATTAATTAATTAATTAATTAATTAAT 220
QY 160 EAOTTPAATEAOTTPGTGLEAOTTAAPAMEAOTTAAPAMEAOTTPPAAMEAOTTPAAT 219
DB 221 TAATTTAATTAATTAATTAATTAATTSATTAATTAATTAATTAATTAATTAATTAAT 280
QY 220 EAOTTAPEATEAOTTPAATEAOTTPLAAMEALSTE-PSATEALSMEPTT 268
DB 281 AATTTAATTAATTAATTAATTTGSPSGSTSTGASTSPSASTATSAPTS 330

```

RESULT 9

Q8VOM1

```

ID Q8VOM1 PRELIMINARY; PRT; 332 AA.
AC Q8VOM1;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
DE Glycoprotein gp2 (Fragment).
OS Equine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae;
NCBI_TaxID=10326;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=1074/94;
RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;
RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034641; AAK61481.1; -
DR InterPro; IPR000104; Antifreeze_1.
DR PRINTS; PR00308; ANTIFREEZE1.
FT NON TER 332
SQ SEQUENCE 332 AA; 30827 MW; 4FB521CD2D525793 CRC64;

```

Query Match 12.6%; Score 256.5; DB 12; Length 332;
 Best Local Similarity 34.7%; Pred. No. 3.6e-08;
 Matches 78; Conservative 9; Mismatches 123; Indels 15; Gaps 2;

```

QY 59 PPEMLNSTDTPLTGTGCTPESTTVEBPARRSTGLDAGAVTELTTELAMNGNSTDSAA 118
DB 101 PTTSTSTTTTTPASTTTPTTTTAAPTTAATTAATTAATTAATTAATTAATTAAT 160
QY 119 MEIOQT-----QPAATEAOTTPLAATEAOTRLTAATEAOTTPLAATEAOTTPAAT 169
DB 161 TPTSTTTTATTTTPTASTTTTDTTAAATTAATTAATTAATTAATTAATTAATTAAT 220
QY 170 EAOTTPGTGLEAOTTAAPAMEAOTTAAPAMEAOTTPPAAMEAOTTPPAAMEAOTTPAAT 229
DB 221 TAATTTAATTAATTAATTAATTSATTAATTAATTAATTAATTAATTAATTAATTAAT 280
QY 230 EAOTTPAATEAOTTPLAAMEALSTE-----EPSATEALSMEPTT 268
DB 281 AATTTAATTAATTAATTTGSPSGSTSTGASTSPSASTATSAPTS 325

```

RESULT 10

```

O39307 PRELIMINARY; PRT; 750 AA.
ID O39307
AC O39307;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE Positional counterpart of HSV-1 gene US5.
OS Equine herpesvirus 4.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae;
NCBI_TaxID=10331;
RN NCB1
RP SEQUENCE OF 685-750 FROM N.A.
RC STRAIN=NS80567;
RA MEDLINE=93189454; PubMed=8397286;
RA Cullinan A.A., Neilan J., Wilson L., Davison A.J., Allen G.;
RT "The DNA sequence of the equine herpesvirus 4 gene encoding glycoprotein gp17/18, the homologue of herpes simplex virus glycoprotein gp.",
RL J. Gen. Virol. 74:1959-1964 (1993).

```

```

RP SEQUENCE OF 1-111 FROM N.A.
RC STRAIN=NS80567;
RX MEDLINE=93119267; PubMed=8380320;
RA Nagesha H.S., Crabb B.S., Studdert M.J.;
RT "Analysis of the nucleotide sequence of five genes at the left end of the unique short region of the equine herpesvirus 4 genome."

```


Db 767 T-----TTAETSTTEPSSSNTPV 786

RESULT 13

OSVOM1 PRELIMINARY; PRT; 293 AA.

ID OSVOM1

AC OSVOM1; 01-MAR-2002 (TEMBLrel. 20, Created)

DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE Glycoprotein gp2 (Fragment).

OS Equine herpesvirus 4.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

OC NCB1_TaxID=10331;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=R19/68;

RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;

RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."

RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY034660; AAKS8439.1; -

DR InterPro: IPR000104; Antifreeze_1.

DR PRINTS; PR00308; ANTIFREEZE1.

FT NON TER 1

FT NON TER 293

SQ SEQUENCE 293 AA; 27805 MW; 7432FEAD5EBE9FF CRC64;

Query Match 12.0%; Score 244; DB 12; Length 293;

Best Local Similarity 37.4%; Pred. No. 1.8e-07;

Matches 80; Conservative 11; Mismatches 111; Indels 12; Gaps 4;

QY 57 TEPPPMRNSTDTPLTGPCTPESTVEPARRSTGIDAGAVTELTELAMNGNSTDS 116

DB 3 TSVSEPTSTATTAT--TTTST--ATTSTTATTTATTTSTTTSTTA 57

QY 117 AAMEIOTTPATEAOTTPLAATEAOTRLATEAOTTPLAATEAOTTPATEAOTTP 176

DB 58 ATTSTSTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 115

QY 177 TGLEAOTTPAMEAOTTPAMEAOTTPAMEAOTTPAMEAOTTPAMEAOTTP 236

DB 116 --TAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 173

QY 237 TATEAOT--TPLAMENSTEPSATLSMEPT 267

DB 174 TTTAATTTATPTESSEAYSTLAATTADTTDTT 207

RESULT 14

OSVOM4 PRELIMINARY; PRT; 316 AA.

ID OSVOM4

AC OSVOM4; 01-MAR-2002 (TEMBLrel. 20, Created)

DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE Glycoprotein gp2 (Fragment).

OS Equine herpesvirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

OC NCB1_TaxID=10326;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=848/89;

RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;

RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."

RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY034636; AAK61476.1; -

DR InterPro: IPR000104; Antifreeze_1.

DR PRINTS; PR00308; ANTIFREEZE1.

FT NON TER 316

SQ SEQUENCE 316 AA; 29361 MW; F8443BE275E06DA8 CRC64;

Query Match 12.0%; Score 243.5; DB 12; Length 316;

Best Local Similarity 35.5%; Pred. No. 2e-07;

Matches 75; Conservative 13; Mismatches 120; Indels 3; Gaps 3;

QY 59 PEMLRNSTDTPLTGPCTPESTVEPARRSTGIDAGAVTELTELAMNGNSTDS 118

DB 101 PSTSAETTTTPTASTTTPTTTTAAPTAAATTTATTAATTAATTAATTAATTA 160

QY 119 MEIOTTPATEAOTTPLAATEAOTRLATEAOTTPLAATEAOTTPATEAOTTP 178

DB 161 TPTSTTTTATTTVPT-TASTTTDTTTATTTAAATTTAAATTTAAATTTAA 219

QY 179 LEAOTTPAMEAOTTPAMEAOTTPAMEAOTTPAMEAOTTPAMEAOTTPATE 238

DB 220 TTAATTS-SATTAATTSATTAATTTATTTATTTATTTATTTATTTATTTA 278

QY 239 TEAOTTPLAAMEALSTE-PSATEALSMEPT 268

DB 279 TGSPTSGSTTGTGASTSTPSASTSATPTS 309

RESULT 15

OSVOM0 PRELIMINARY; PRT; 389 AA.

ID OSVOM0

AC OSVOM0; 01-MAR-2002 (TEMBLrel. 20, Created)

DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)

DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE Glycoprotein gp2 (Fragment).

OS Equine herpesvirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Varicellovirus.

OC NCB1_TaxID=10326;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1070/94;

RA Huang J.-A., Ficorilli N., Hartley C.A., Allen G.P., Studdert M.J.;

RT "Polymorphism of open reading frame 71 of equine herpesvirus 4 (EHV4) and EHV1."

RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY034640; AAK61480.1; -

DR InterPro: IPR000104; Antifreeze_1.

DR PRINTS; PR00308; ANTIFREEZE1.

FT NON TER 389

FT NON TER 389

SQ SEQUENCE 389 AA; 42357630A8C3A592 CRC64;

Query Match 11.9%; Score 241; DB 12; Length 389;

Best Local Similarity 35.7%; Pred. No. 3.6e-07;

Matches 82; Conservative 14; Mismatches 114; Indels 20; Gaps 5;

QY 55 PETPEPMRNSTDTPLTGPCTPESTVEPARRSTGIDAGAVTELTELAMNGNST 114

DB 157 PTTTPTSTTTTAT--TVPTTASTTTDTTTAAATTTATTTATTTATTTATTT 214

QY 115 DSAAMEIOTTPAAEAOITPLAATEAOTRL-----ATEAOTTPA-----AT 159

DB 215 TTAATTTAATTSATTAATTSATTAATTSATTAATTSATTAATTSATTAATTS 274

QY 160 EAOTTPATEAOTTPGLEAOTTPAMEAOTTPAMEAOTTPAMEAOTTPAMEAOTTP 219

DB 275 TTAATTSATTAATTSAT--TAATTSATTAATTTATTTATTTATTTATTTA 332

QY 220 EAOTTPATEAOTTPATEAOTTPAMEALSTE-PSATEALSMEPT 268

DB 333 AATTTAATTTAATTTAATTTGPTSGSTTGTGASTSTPSASTSATPTS 382

Search completed: July 14, 2003, 06:14:41

Job time : 83 secs


```

FT Region 311..332
FT /label= Transmembrane_region
FT Domain 333..402
FT /label= Cytoplasmic_domain
XX
PN W09530001-A2.
PD 09-NOV-1995.
XX
PF 24-APR-1995; 95WO-US04968.
XX
PR 30-SEP-1994; 94US-0316305.
PR 28-APR-1994; 94US-0235398.
XX
PA (GEMV ) GENETICS INST INC.
XX
PI Chang X, Cumming D, Kumar R, Larsen GR, Sako DS;
PI Shaw GD, Veldman GM;
XX
DR WPI; 1995-403865/51.
DR N-PSDB; AAT02484.
XX
PT Novel P-selectin ligand protein - useful as anti-inflammatory agent
PT and to identify inhibitors of selectin-mediated intercellular
PT adhesion
XX
PS Claim 1; Page 82-83; 135pp; English.
XX
CC P-selectin ligand (AAR85075) is an inhibitor of E- and P-selectin-
CC mediated intercellular adhesion. It was obtd. by expression
CC of cDNA clone PMT21:PH85 (ATCC 69096) (AAT02484), isolated from human
CC promyelocytic HL-60 cells. Recombinant P-selectin ligand can
CC be produced in mammalian host cells co-transfected with a
CC fucosyltransferase gene, for use as an antiinflammatory.
XX
SQ Sequence 402 AA;
XX
Query Match 100.0%; Score 2030; DB 16; Length 402;
Best Local Similarity 100.0%; Pred. No. 2,3e-149;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRROATEVEYLDYDLPTETPP 60
DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRROATEVEYLDYDLPTETPP 60
XX
QY 61 EMLRNSIDTTLPLGPGTPESTVEPAARRSTGLDAGAVTELTELAMNGSLTDSAME 120
DB 61 EMLRNSIDTTLPLGPGTPESTVEPAARRSTGLDAGAVTELTELAMNGSLTDSAME 120
XX
QY 121 IQTTPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGLE 180
DB 121 IQTTPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGLE 180
XX
QY 121 IQTTPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGLE 180
DB 121 IQTTPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGLE 180
XX
QY 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPAATEAQTTPPAE 240
DB 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPAATEAQTTPPAE 240
XX
QY 241 AQTTPAAMEALSTPEATEALSMETTKRGLFPESVSVTHKGIPIMAASNLVAVPVG 300
DB 241 AQTTPAAMEALSTPEATEALSMETTKRGLFPESVSVTHKGIPIMAASNLVAVPVG 300
XX
QY 301 APDHISVKOCLAILILAVATTFVCTVVLAVRLSRKGMYPVRNVSPTEMVCISLAP 360
DB 301 APDHISVKOCLAILILAVATTFVCTVVLAVRLSRKGMYPVRNVSPTEMVCISLAP 360
XX
QY 361 DGEGSGSATNGLSKAKSPGLTPPEPREDGGDLTLHSFLP 402
DB 361 DGEGSGSATNGLSKAKSPGLTPPEPREDGGDLTLHSFLP 402
XX
RESULT 2
AAW26174
ID AAW26174 standard; protein; 402 AA.

```

```

XX AC AAW26174;
XX DT 28-NOV-1997 (first entry)
XX DE P-selectin glycoprotein ligand-1.
XX
XX P-selectin glycoprotein ligand-1; PSG-1, human; mucin-like protein;
XX calcium-dependent carbohydrate binding protein; platelet; endothelium;
XX thrombin; leukocyte; sulphated glycosylated peptide; O-glycan; therapy;
XX inflammation; ischaemia; reperfusion; bacterial sepsis; atherosclerosis;
XX disseminated intravascular coagulation; rheumatoid arthritis; antibody;
XX adult respiratory distress syndrome; tumour metastasis;
XX autoimmune disorder; inflammatory disorder.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..19 /note= "signal peptide"
XX Peptide 19..41 /note= "propeptide"
XX Protein 42..402 /note= "mature PSG-1"
XX Domain 42..308
XX Domain /note= "extracellular domain"
XX Domain 309..333 /note= "transmembrane domain"
XX Domain 334..402 /note= "cytoplasmic tail"
XX
XX W09706176-A2.
XX
XX 20-FEB-1997.
XX
XX 02-AUG-1996; 96WO-US12820.
XX
XX 17-MAY-1996; 96US-0649802.
XX 03-AUG-1995; 95US-0510920.
XX 15-MAY-1996; 96US-0017794.
XX
XX (OKLA ) UNIV OKLAHOMA STATE.
XX
XX Cummings RD, Mcever RP, Moore KL;
XX
XX WPI; 1997-154206/14.
XX
XX New O-glycan(e) derived from P-selectin glyco-protein ligand
XX (PSGL)-1 - used for inhibiting binding of PSG-1 to selectins, e.g.
XX for treating inflammatory or autoimmune disorders or tumours
XX
XX Claim 16; Page 78-79; 96pp; English.
XX
XX This sequence represents the human P-selectin glycoprotein ligand-1
XX (PSGL-1). P-selectin is a calcium-dependent carbohydrate binding protein
XX expressed on the surfaces of activated platelets and endothelium in
XX response to thrombin and other agonists. PSG-1 is a high affinity
XX P-selectin ligand produced by leukocytes. Binding of P-selectin to
XX PSG-1 is calcium ion dependent and is abolished by treatment of the
XX ligand with sialidase. PSG-1 is a homodimer, with a highly extended
XX extracellular domain, which is a feature of mucin-like proteins. This
XX sequence is targeted by the sulphated glycosylated peptides of the
XX invention (see AAW26175-W26188). PSG-1 is also targeted by the
XX O-glycans, and O-glycanated products of the invention. The products are
XX used for inhibiting the binding of P-selectin and other selectins to
XX PSG-1. They can be used for the treatment of inflammation, injury
XX resulting from ischaemia and reperfusion, bacterial sepsis and
XX disseminated intravascular coagulation, adult respiratory distress
XX syndrome, tumour metastasis, rheumatoid arthritis, atherosclerosis and
XX other autoimmune or inflammatory disorders. The products can also be used
XX in the production of antibodies and in diagnostic applications.
XX
XX Sequence 402 AA;

```


Query Match 100.0%; Score 2030; DB 18; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLQILLILILGPNLSIQMTWADAEKALGPIIARDRQATEYEDYFLPETEP 60
 DB 1 MFLQILLILILGPNLSIQMTWADAEKALGPIIARDRQATEYEDYFLPETEP 60
 QY 61 EMLRSTDTPLTGPGTPESTVEPARRSTGLDAGAVTELTTELANKGNLSTSAAME 120
 DB 61 EMLRSTDTPLTGPGTPESTVEPARRSTGLDAGAVTELTTELANKGNLSTSAAME 120
 QY 121 IOTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGIE 180
 DB 121 IOTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGIE 180
 QY 181 AOTTAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPATE 240
 DB 181 AOTTAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPATE 240
 QY 241 AOTTPLAMEALSTPSATEALSMETTRKGLFIPFSVSVTHKGIIPMAASNLVNYPVG 300
 DB 241 AOTTPLAMEALSTPSATEALSMETTRKGLFIPFSVSVTHKGIIPMAASNLVNYPVG 300
 QY 301 APDHISVQCLAILILALVATIFVCTVLAVALRSRKGHPVANYSPTEWVCISSLP 360
 DB 301 APDHISVQCLAILILALVATIFVCTVLAVALRSRKGHPVANYSPTEWVCISSLP 360
 QY 361 DGGEGPSATANGSLSKAKSPGLTPPEPRDREGDILLHSFLP 402
 DB 361 DGGEGPSATANGSLSKAKSPGLTPPEPRDREGDILLHSFLP 402

RESULT 3
 AAW72695
 ID AAW72695 standard; Protein; 402 AA.
 AC AAW72695;
 XX
 DT 08-JAN-1999 (first entry)
 XX
 DE Human glycoprotein P-selectin ligand protein.
 XX
 KM Human; glycoprotein; P-selectin ligand protein; inflammatory disease;
 XX intracellular domain; E-selectin-mediated intercellular adhesion.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 42..60 /note= "see claim 2"
 FT Misc-difference 42..88 /note= "see claim 3"
 FT Misc-difference 42..118 /note= "see claim 5"
 FT Misc-difference 42..189 /note= "see claim 6"
 FT Misc-difference 42..310 /note= "see claim 8"
 FT Misc-difference 42..316 /note= "see claim 10"
 FT
 FT
 PN US5827817-A.
 XX
 PD 27-OCT-1998.
 XX
 PF 07-JUN-1995; 95US-0477254.
 XX
 PR 25-APR-1995; 95US-0428734.
 PR 23-OCT-1992; 92US-0965662.
 PR 26-AUG-1993; 93US-0112608.
 PR 22-OCT-1993; 93WO-US10168.

PR 28-APR-1994; 94US-0235398.
 PR 30-SEP-1994; 94US-0316305.
 PR 07-JUN-1995; 95US-0477254.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Camphausen R, Chang X, Cumming D, Kumar R, Larsen GR;
 PI Sako DS, Shaw G, Veldman GM;
 XX
 DR WPI: 1998-594030/50.
 DR N-PSDB; AAW67130.
 XX
 PT Soluble P-selectin ligand glycoprotein - with truncated protein
 PT sequence lacking intracellular domain
 XX
 PS Claim 1; Column 47-50; 66pp; English.

CC The present sequence represents a human glycoprotein having P-selectin
 CC ligand (PSL) activity. The present invention describes a glycoprotein
 CC having PSL activity which comprises a carbohydrate and a protein. The
 CC protein includes amino acids 42-60 of AAW72695, but lacks the
 CC intracellular domain and optionally the transmembrane domain of mature
 CC PSL protein, and the carbohydrate contains sialyl Lewis X. The isolated
 CC PSL protein may be useful in treating conditions characterized by P- or
 CC E-selectin mediated intercellular adhesion. Such conditions include
 CC myocardial infarction, bacterial or viral infection, metastatic
 CC conditions, inflammatory disorders, asthma, emphysema, thermal injury,
 CC multiple sclerosis, diabetes, Reynaud's syndrome, neutrophilic
 CC dermatosis, Grave's disease, glomerulonephritis, gingivitis,
 CC periodontitis, hemolytic uremic syndrome, ulcerative colitis, Crohn's
 CC disease, necrotizing enterocolitis, and cytokine-induced toxicity. The
 CC isolated PSL protein may also be useful in organ transplantation, both to
 CC prepare organs for transplantation and to quell organ transplant
 CC rejection. The isolated PSL may be used to treat haemodialysis and
 CC leukophoresis patients. Additionally, isolated PSL protein may be used as
 CC an antimetastatic agent. The isolated PSL protein may be used itself as
 CC an inhibitor of P- or E-selectin-mediated intercellular adhesion or to
 CC design inhibitors of P- or E-selectin-mediated intercellular adhesion.
 CC
 XX

Sequence 402 AA;
 SQ

Query Match 100.0%; Score 2030; DB 19; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLQILLILILGPNLSIQMTWADAEKALGPIIARDRQATEYEDYFLPETEP 60
 DB 1 MFLQILLILILGPNLSIQMTWADAEKALGPIIARDRQATEYEDYFLPETEP 60
 QY 61 EMLRSTDTPLTGPGTPESTVEPARRSTGLDAGAVTELTTELANKGNLSTSAAME 120
 DB 61 EMLRSTDTPLTGPGTPESTVEPARRSTGLDAGAVTELTTELANKGNLSTSAAME 120
 QY 121 IOTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGIE 180
 DB 121 IOTTQPAATEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPGIE 180
 QY 181 AOTTAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPATE 240
 DB 181 AOTTAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPAPAMEAQTTPATE 240
 QY 241 AOTTPLAMEALSTPSATEALSMETTRKGLFIPFSVSVTHKGIIPMAASNLVNYPVG 300
 DB 241 AOTTPLAMEALSTPSATEALSMETTRKGLFIPFSVSVTHKGIIPMAASNLVNYPVG 300
 QY 301 APDHISVQCLAILILALVATIFVCTVLAVALRSRKGHPVANYSPTEWVCISSLP 360
 DB 301 APDHISVQCLAILILALVATIFVCTVLAVALRSRKGHPVANYSPTEWVCISSLP 360
 QY 361 DGGEGPSATANGSLSKAKSPGLTPPEPRDREGDILLHSFLP 402
 DB 361 DGGEGPSATANGSLSKAKSPGLTPPEPRDREGDILLHSFLP 402

ID	AAW53320	standard; Protein; 402 AA.
AC	AAW53320;	
XX	27-JUL-1998	(first entry)
DE	P-selectin ligand amino acid sequence.	
XX		
KM	P-selectin ligand protein; glycoprotein; human; calcium-dependent;	
KW	intercellular adhesion; E-selectin; L-selectin; myocardial infarction	
KW	inflammation; Crohn's disease; transplant rejection; rejection1.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..41
FT		/note= "signal peptide"
FT	Protein	42..402
FT		/note= "mature protein"
FT	Protein	21..310
FT		/note= "soluble for of the protein"
FT	Protein	42..310
FT		/note= "soluble form of the protein"
FT	Cleavage-site	38..41
FT		/note= "PACE (Paired basic Amino acid Converting Enzyme cleavage site"
FT	Domain	21..310
FT		/note= "extracellular domain"
FT	Domain	311..332
FT		/note= "transmembrane domain "
FT	Domain	333..402
FT		/note= "intracellular, cytoplasmic domain"
FT	Modified-site	65..67
FT		/note= "Asn is potentially glycosylated"
FT	Modified-site	111..113
FT		/note= "Asn is potentially glycosylated"
FT	Modified-site	292..294
FT		/note= "Asn is potentially glycosylated"
FT	Modified-site	46
FT		/note= "potential tyrosine sulfation site"
FT	Modified-site	48
FT		/note= "potential tyrosine sulfation site"
FT	Modified-site	51
FT		/note= "potential tyrosine sulfation site"
XX		
FN	W09808949-A1.	
PD	05-MAR-1998.	
XX		
XX	29-AUG-1997;	97MO-US14159.
XX		
PR	30-AUG-1996;	96US-0713556.
XX		
PA	(GENMY) GENETICS INST INC.	
XX		
PI	Chang X, Cumming D, Kumar R, Larsen GR, Sako DS;	
PI	Shaw G, Veldman GM;	
XX		
DR	WPI; 1998-179434/16.	
DR	N-PSDB; AAV20880.	
XX		
PT	New isolated DNA encoding fusion protein including P-selectin ligand	
PT	fragment - to direct second component, e.g. cytokine, to sites of	
PT	formation expression, used e.g. to stimulate bone and cartilage	
XX		
PS	Claim 30; Pages 64-65; 128pp; English.	
CC	The present sequence represents a P-selectin ligand protein. The	
CC	P-selectin ligand is a glycoprotein. The extracellular domain of the	

Query	Match	Similarity	Score	DB	Length	402;
Best Local	Similarity	100.0%;	Score 2030;	DB 19;	Length 402;	
Matches	402;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0
QY	1	MPQLLLILLIGGNSIQLMWDVADEKALGPLAADRRQATYEVYDYPFLPETEP	60			
DB	1	MPQLLLILLIGGNSIQLMWDVADEKALGPLAADRRQATYEVYDYPFLPETEP	60			
QY	61	EMLNSTDTPLTGPGETSTTVEPAARSTGLDAGAVTELTLTLANNGNSTPSAME	120			
DB	61	EMLNSTDTPLTGPGETSTTVEPAARSTGLDAGAVTELTLTLANNGNSTPSAME	120			
QY	121	IQTTOPATEAQTPPLATEAQTRRLTBEAQTPPLATEAQTPPAAATEAQTPPTGLE	180			
DB	121	IQTTOPATEAQTPPLATEAQTRRLTBEAQTPPLATEAQTPPAAATEAQTPPTGLE	180			
QY	181	AQTTAPPAAMEAQTTAPPAAMEAQTTTPPAMEAQTTTPPAAATEAQTPPTATE	240			
DB	181	AQTTAPPAAMEAQTTAPPAAMEAQTTTPPAMEAQTTTPPAAATEAQTPPTATE	240			
QY	241	AQTPPLAMEALSTEPSATEALNSEPTKRGFTFVSYSVTHKGI PMAASLVNYPVG	300			
DB	241	AQTPPLAMEALSTEPSATEALNSEPTKRGFTFVSYSVTHKGI PMAASLVNYPVG	300			
QY	301	APDHISVQCLALILIALVATIFPVCTVLAVALRSRGHMPVANYSPTEWVCISLLP	360			
DB	301	APDHISVQCLALILIALVATIFPVCTVLAVALRSRGHMPVANYSPTEWVCISLLP	360			
QY	361	DGEGSPSATANGCISLKAKSPGLTPEPRDRBGDDITLHSFLP	402			
DB	361	DGEGSPSATANGCISLKAKSPGLTPEPRDRBGDDITLHSFLP	402			
RESULT 5						
AA29765						
ID	AA29765 standard; Protein; 402 AA.					
AC	AA29765;					
XX	04-NOV-1999 (first entry)					
XX	Human P-selectin ligand protein.					
KW	Human; P-selectin ligand; glycoprotein; fusion protein; infection;					
KW	inflammation; intercellular adhesion; ulcerative colitis; asthma;					
KW	diabetes; transplant rejection; myocardial infarction; thermal injury;					
KW	metabolic condition; autoimmune thyroiditis; multiple sclerosis;					
KW	Reynaud's syndrome; neutrophilic dermatosis; Sweet's syndrome;					
KW	Graves's disease; glomerulonephritis; gingivitis; periodontitis;					
KW	Crohn's disease; necrotising enterocolitis.					
DS	Homo sapiens					


```

FT Modified-site 65..67
FT /note= "Asn is N-glycosylated"
FT Modified-site 111..113
FT /note= "Asn is N-glycosylated"
FT Modified-site 292..294
FT /note= "Asn is N-glycosylated"
FT Domain
FT 311..332
FT /label= Transmembrane_domain
FT 333..402
FT /label= Intracellular_cytoplasmic_domain
PN WO200172769-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US09815.
XX
XX 27-MAR-2000; 2000US-193351P.
XX
XX (GEMV ) GENETICS INST INC.
XX
XX Coffman JL, Foster WB, Germain BJ, Sun S, Robinson JJ;
XX WPI; 2001-616464/71.
XX
XX Purifying highly anionic target proteins from impurities and
XX DNA/histone complexes in sample, by contacting sample to substrate that
XX binds target, while impurities are washed off and the complexes are
XX dissociated -
XX
XX Disclosure; Page 30-31; 31pp; English.
XX
XX The present invention provides a method for purifying highly anionic
XX target proteins from proteinaceous and non-proteinaceous impurities
XX and DNA/histone complexes in the sample. The method of the invention
XX involves contacting the sample with a substrate that reversibly binds
XX charged molecules, and purifying, under conditions such that the
XX impurities are not bound to or washed off the substrate. The method is
XX then dissociated, while the target protein remains bound. The method is
XX useful for isolating and purifying highly anionic target proteins e.g.
XX sulphated proteins, from proteinaceous and non-proteinaceous impurities,
XX and DNA/histone complexes in the sample. The present sequence is of
XX the human P-selectin glycoprotein ligand-1 (PSGL-1) protein used in the
XX invention. The ten amino acid consensus repeat of the decameric
XX subdomain of this protein is indicated in AAU09942.
XX
XX
XX Sequence 402 AA;
XX
XX Query Match 100.0%; Score 2030; DB 22; Length 402;
XX Best Local Similarity 100.0%; Pred. No. 2,3e-149;
XX Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPQLQLLLILILGPGNSLQIMDTWADBAKALGPLIARDRRQATEYEDYDFLPETPP 60
DB 1 MPQLQLLLILILGPGNSLQIMDTWADBAKALGPLIARDRRQATEYEDYDFLPETPP 60
QY 61 EMLRNSIDTTLPLGPGSTTTEPARBSTGLDAGGAVTELTTLANNGNSTDSAAAE 120
DB 61 EMLRNSIDTTLPLGPGSTTTEPARBSTGLDAGGAVTELTTLANNGNSTDSAAAE 120
QY 121 IOTTOPAAEAQTTPLAATEAQTTRJLATEAQTTPLAATEAQTTPPAATEAQTTPGHE 180
DB 121 IOTTOPAAEAQTTPLAATEAQTTRJLATEAQTTPLAATEAQTTPPAATEAQTTPGHE 180
QY 122 IOTTOPAAEAQTTPLAATEAQTTRJLATEAQTTPLAATEAQTTPPAATEAQTTPGHE 180
DB 122 IOTTOPAAEAQTTPLAATEAQTTRJLATEAQTTPLAATEAQTTPPAATEAQTTPGHE 180
QY 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAE 240
DB 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAE 240
QY 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAE 300
DB 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAE 300
QY 301 APDHISVKOCLLAILLALVATIFVCTVVLAVLRKGMYPVRYNSPTENVCISSLLP 360

```

```

DB 301 APDHISVKOCLLAILLALVATIFVCTVVLAVLRKGMYPVRYNSPTENVCISSLLP 360
QY 361 DGEGPSATANGGLSKAKSPGLTPPEPREDEDDTLHSFLP 402
DB 361 DGEGPSATANGGLSKAKSPGLTPPEPREDEDDTLHSFLP 402
RESULT 9
AAU79653
ID AAU79653 standard; Protein; 402 AA.
XX
XX AAU79653;
XX
XX 02-JUL-2002 (first entry)
XX
XX Human P-selectin glycoprotein ligand-1 (PSGL-1).
XX
XX Human; stenosis; restenosis; vascular injury; cardiovascular disease;
XX soluble P-selectin glycoprotein ligand-1; PSGL-1; vascular remodeling;
XX neointimal formation; coronary artery disease; peripheral artery;
XX carotid artery; percutaneous transluminal coronary angioplasty; PTCA;
XX stent implantation; leukocyte recruitment; cell to cell adhesion;
XX blood vessel; vasotropic; cardiac.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..41
XX /label= Signal_peptide
XX Protein 42..402
XX /label= Mature_PSGL-1
XX Region 42..316
XX /note= "This region is specifically claimed in claim 24"
XX
XX WO200222820-A1.
XX
XX 21-MAR-2002.
XX
XX 12-SEP-2000; 2000WO-US25007.
XX
XX 12-SEP-2000; 2000WO-US25007.
XX
XX (GEMV ) GENETICS INST INC.
XX (MONT-) MONTREAL HEART INST.
XX
XX Kumar A, Schnab RG, Tanguay J, Werhi Y;
XX WPI; 2002-351889/38.
XX N-PSDB; ABR48715.
XX
XX Modulating stenosis or restenosis in subject having vascular injury or
XX cardiovascular disease, comprises administering P-selectin antagonist
XX to the subject -
XX
XX Claim 24; Page 57-59; 66pp; English.
XX
XX The present invention relates to methods and compositions for modulating
XX stenosis or restenosis in a subject having vascular injury or
XX cardiovascular disease. The method comprises administering a P-selectin
XX antagonist such as soluble P-selectin glycoprotein ligand-1 (PSGL-1) to
XX the subject. The method is useful for preventing, inhibiting or treating
XX stenosis or restenosis in a subject, preferably human, having vascular
XX injury or coronary artery disease. Restenosis is characterized by
XX constrictive vascular remodeling or neointimal formation. The vascular
XX injury or cardiovascular disease affects a coronary artery, or peripheral
XX artery, preferably a carotid artery. The vascular injury results from
XX angioplasty (percutaneous transluminal coronary angioplasty (PTCA)), or
XX implantation of a stent or stents. The method is also useful for
XX modulating leukocyte recruitment, inhibiting cell to cell adhesion, or
XX inhibiting cell adhesion to blood vessels in a subject. The present
XX sequence represents human PSGL-1.
XX

```

SQ Sequence 402 AA;
 Query Match 100.0%; Score 2030; DB 23; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEDYDFLPETEP 60
 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEDYDFLPETEP 60
 DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEDYDFLPETEP 60
 QY 61 EMLRNSDTPTPLGPGTPESTVPEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 61 EMLRNSDTPTPLGPGTPESTVPEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 DB 61 EMLRNSDTPTPLGPGTPESTVPEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 QY 121 IOTTOPAAEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPG 180
 121 IOTTOPAAEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPG 180
 DB 121 IOTTOPAAEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPG 180
 QY 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 240
 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 240
 DB 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 240
 QY 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 300
 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 300
 DB 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 300
 QY 301 APDHISVKOCCLAILLALVATTFVCTVLAVALSRKGMYPVARNYSTEWVCISILP 360
 301 APDHISVKOCCLAILLALVATTFVCTVLAVALSRKGMYPVARNYSTEWVCISILP 360
 DB 301 APDHISVKOCCLAILLALVATTFVCTVLAVALSRKGMYPVARNYSTEWVCISILP 360
 QY 361 DGEGPSATANGLSKAKSPGLTPPREDEGDDLTLSFLP 402
 361 DGEGPSATANGLSKAKSPGLTPPREDEGDDLTLSFLP 402
 DB 361 DGEGPSATANGLSKAKSPGLTPPREDEGDDLTLSFLP 402

RESULT 10
 AAG77945
 ID AAG77945 standard; Protein; 402 AA.
 AC AAG77945;
 XX
 DT 31-JAN-2002 (first entry)
 XX
 DE Human P-selectin glycoprotein ligand-1.
 XX
 KM Human; PSGL-1; P-selectin glycoprotein ligand-1; thrombosis;
 KM thrombolytic; anticoagulant; cardiovascular disease; atherosclerosis;
 KM hypertension; angioplasty.
 XX
 OS Homo sapiens.
 XX
 PN MO200175107-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 02-APR-2001; 2001MO-US10622.
 XX
 PR 31-MAR-2000; 2000US-193787P.
 XX
 PA (GEM) GENETICS INST INC.
 XX
 PI Eppihimer MJ, Schaub RG, Harris AS;
 XX
 DR WPI; 2002-010790/01.
 DR N-PSDB; AAH77236.
 XX
 PT Modulating (e.g. preventing, inhibiting or treating) thrombosis,
 PT comprises administering a P-selectin antagonist to subject, e.g. human
 PT with cardiovascular disorder (e.g. atherosclerosis or hypertension) -
 XX
 PS Claim 13; Page 69-70; 73pp; English.
 XX
 CC The sequence represents human P-selectin glycoprotein ligand-1 (PSGL-1).

CC The invention relates to a novel method for treating or inhibiting
 CC thrombosis in a subject comprising administering a composition comprising
 CC an effective amount of a P-selectin antagonist. The polypeptides and
 CC polynucleotides of the invention have thrombolytic, and anticoagulant
 CC activity. The method is useful for treating or inhibiting thrombosis in a
 CC subject, particularly a human. The method is also useful for inhibiting
 CC cell adhesion to blood vessels in a subject, increasing the movement of
 CC cells relative to blood vessels in a subject, or inhibiting the effect of
 CC a thrombus-inducing agent in a subject. The subject is at risk for
 CC thrombosis, e.g. a subject suffering from a cardiovascular disease or
 CC disorder (e.g. atherosclerosis or hypertension), or a subject who has
 CC undergone cardiovascular or general vascular procedures or intervention
 CC such as angioplasty of any vessel.
 XX

SQ Sequence 402 AA;
 Query Match 100.0%; Score 2030; DB 23; Length 402;
 Best Local Similarity 100.0%; Pred. No. 2,3e-149;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEDYDFLPETEP 60
 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEDYDFLPETEP 60
 DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEDYDFLPETEP 60
 QY 61 EMLRNSDTPTPLGPGTPESTVPEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 61 EMLRNSDTPTPLGPGTPESTVPEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 DB 61 EMLRNSDTPTPLGPGTPESTVPEPARRSTGLDAGAVTELTTLANMGNLSTDSAME 120
 QY 121 IOTTOPAAEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPG 180
 121 IOTTOPAAEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPG 180
 DB 121 IOTTOPAAEAQTTPLAATEAQTTRLTATEAQTTPLAATEAQTTPPAATEAQTTPG 180
 QY 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 240
 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 240
 DB 181 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 240
 QY 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 300
 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 300
 DB 241 AQTTPAAMEAQTTPAAMEAQTTPPAAMEAQTTPPAAMEAQTTPPAATEAQTTPG 300
 QY 301 APDHISVKOCCLAILLALVATTFVCTVLAVALSRKGMYPVARNYSTEWVCISILP 360
 301 APDHISVKOCCLAILLALVATTFVCTVLAVALSRKGMYPVARNYSTEWVCISILP 360
 DB 301 APDHISVKOCCLAILLALVATTFVCTVLAVALSRKGMYPVARNYSTEWVCISILP 360
 QY 361 DGEGPSATANGLSKAKSPGLTPPREDEGDDLTLSFLP 402
 361 DGEGPSATANGLSKAKSPGLTPPREDEGDDLTLSFLP 402
 DB 361 DGEGPSATANGLSKAKSPGLTPPREDEGDDLTLSFLP 402

RESULT 11
 AAR53965
 ID AAR53965 standard; Protein; 402 AA.
 AC AAR53965;
 XX
 DT 14-OCT-1994 (first entry)
 XX
 DE P-selectin ligand.
 XX
 KM P-selectin ligand; glycoprotein; inflammation;
 KM intercellular adhesion; promyelocyte; HL60; PM721; PL85;
 KM alpha-1,3/alpha-1,4-fucosyltransferase; PACSOL;
 KM paired basic amino acid converting enzyme.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Peptide 1..41
 FT /label= Sig_peptide
 FT Domain 21..310
 FT /label= extracellular domain
 FT Misc-difference 46
 FT /note= "potential sulfation site"

```

FT Misc-difference 48 /note= "potential sulfation site"
FT Misc-difference 51 /note= "potential sulfation site"
FT Misc-difference 65 /note= "potential N-linked glycosylation site"
FT Misc-difference 111 /note= "potential N-linked glycosylation site"
FT Misc-difference 292 /note= "potential N-linked glycosylation site"
FT M1sc-difference 292 /note= "potential N-linked glycosylation site"
FT Domain 311..332 /label= "transmembrane domain"
FT Domain 333..402 /label= "cytoplasmic domain"
FT Domain /label= "cytoplasmic domain"
PN MO9410309-A.
XX 11-MAY-1994.
XX 22-OCT-1993; 93WO-US10168.
XX 23-OCT-1992; 92US-0965662.
XX 26-AUG-1993; 93US-0112608.
XX (GENY ) GENETICS INST INC.
XX Chang X, Larsen GR, Sako DS, Veldman GM;
XX WPI: 1994-167466/20.
XX N-PSDB; Q-63953.
XX
XX DNA encoding P-selectin ligand glyco:protein - for use in
XX treating inflammatory diseases characterised by P-selectin
XX mediated inter:cellular adhesion
XX
XX Disclosure; Page 48-50; 64pp; English.
XX
XX cDNA of sequence AAQ63953, encoding P-selectin ligand, was derived
XX from promyelocyte line HL60 clone PMT21:PL85. The ligand comprises
XX residues 1-402, 1-310 (sol.), 42-402 (mature) or 42-310 (sol.).
XX mature) of sequence AAR53965. An allelic variant (AAQ63954, AAR53966)
XX was derived from human placenta. Co-transfection of a mammalian host
XX cell with DNA encoding residues 42-310, plus DNA encoding a soluble
XX alpha-1,3/alpha-1,4-fucosyltransferase and DNA encoding a soluble
XX form of a paired basic amino acid converting enzyme (PACESOL).
XX CC encoded by sequence AAQ63957, allowed production of sol. mature
XX P-selectin ligand glycoprotein.
XX
XX Sequence 402 AA:
SQ
Query Match 99.6%; Score 2022; DB 15; Length 402;
Best Local Similarity 99.5%; Pred. No. 9.5e-149; Indels 0; Gaps 0;
Matches 400; Conservative 1; Mismatches 1;
QY 1 MPQLILLILILPGNSLQIMTWADEAKALGPLARDRQATEYEDYDPLPETEP 60
Db 1 MPQLILLILILPGNSLQIMTWADEAKALGPLARDRQATEYEDYDPLPETEP 60
QY 61 EMKRNSTDTPLTGPSTTTPPARASTGIDAGAVTELTTEIANNGNSTDSAA 120
Db 61 EMKRNSTDTPLTGPSTTTPPARASTGIDAGAVTELTTEIANNGNSTDSAA 120
QY 121 IOTTTPAATEAQTTPAATEAQTTRLTATEAQTTPAATEAQTTPAATEAQTTP 180
Db 121 IOTTTPAATEAQTTPAATEAQTTRLTATEAQTTPAATEAQTTPAATEAQTTP 180
QY 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 240
Db 181 AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 240
QY 241 AQTTPAAMEALSTESATBALSMETTKGLFIPSVSVVTHKGIIPMAASNL 300
Db 241 AQTTPAAMEALSTESATBALSMETTKGLFIPSVSVVTHKGIIPMAASNL 300

```

```

QY 301 ABPHISVKOCLAIILALVATIFPVCTVLAVALSRKGMYPVRYSPTEWYCIS 360
Db 301 ABPHISVKOCLAIILALVATIFPVCTVLAVALSRKGMYPVRYSPTEWYCIS 360
QY 361 DGEGBSATANGLSKAKSPGLTPPREPREDDTLTSHFLP 402
Db 361 DGEGBSATANGLSKAKSPGLTPPREPREDDTLTSHFLP 402
RESULT 12
AAW72696
ID AAW72696 standard; Protein; 412 AA.
XX
XX AAW72696;
XX
XX 08-JAN-1999 (first entry)
XX
XX Human glycoprotein P-selectin ligand protein.
XX
XX Human; glycoprotein; P-selectin ligand protein; inflammatory disease;
XX intracellular domain; E-selectin-mediated intercellular adhesion.
XX Homo sapiens.
XX
XX US5827817-A.
XX
XX 27-OCT-1998.
XX
XX 07-JUN-1995; 95US-0477254.
XX
XX 25-APR-1995; 95US-0428734.
XX 23-OCT-1992; 92US-0965662.
XX 26-AUG-1993; 93US-0112608.
XX 22-OCT-1993; 93WO-US10168.
XX 28-APR-1994; 94US-0235398.
XX 30-SEP-1994; 94US-0316305.
XX 07-JUN-1995; 95US-0477254.
XX
XX (GENY ) GENETICS INST INC.
XX
XX Camphausen R, Chang X, Cumming D, Kumar R, Larsen GR;
XX Sako DS, Shaw G, Veldman GM;
XX WPI: 1998-594030/50.
XX N-PSDB; AAV67131.
XX
XX Soluble P-selectin ligand glycoprotein - with truncated protein
XX sequence lacking intracellular domain
XX
XX Disclosure; Column 49-52; 66pp; English.
XX
XX The present sequence represents a human glycoprotein having P-selectin
XX ligand (PSL) activity. The present invention describes a glycoprotein
XX having PSL activity which comprises a carbohydrate and a protein. The
XX protein includes amino acids 42-60 of AAW72695, but lacks the
XX intracellular domain and optionally the transmembrane domain of mature
XX PSL protein, and the carbohydrate contains sialyl Lewis x. The isolated
XX PSL protein may be useful in treating conditions characterised by P- or
XX E-selectin mediated intercellular adhesion. Such conditions include
XX myocardial infarction, bacterial or viral infection, metastatic
XX conditions, inflammatory disorders, asthma, emphysema, thermal injury,
XX multiple sclerosis, diabetes, Reynaud's syndrome, neutrophilic
XX dermatosis, Grave's disease, glomerulonephritis, gingivitis,
XX periodontitis, haemolytic uremic syndrome, ulcerative colitis, Crohn's
XX disease, necrotising enterocolitis, and cytokine-induced toxicity. The
XX isolated PSL protein may also be useful in organ transplantation, both to
XX prepare organs for transplantation and to quell organ transplant
XX rejection. The isolated PSL may be used to treat hemodialysis and
XX leukopheresis patients. Additionally, isolated PSL protein may be used as
XX an antineoplastic agent. The isolated PSL protein may be used itself as
XX an inhibitor of P- or E-selectin-mediated intercellular adhesion or to
XX design inhibitors of P- or E-selectin-mediated intercellular adhesion.

```


XX SQ Sequence 412 AA;
 Query Match 99.3%; Score 2015; DB 19; Length 412;
 Best Local Similarity 97.6%; Pred. No. 3.4e-148;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEVLDYDFLPETEP 60
 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEVLDYDFLPETEP 60
 DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEVLDYDFLPETEP 60
 QY 61 EMLNSTDTTPLTGPGTPESTVVEPARRSTGLDAGAVTELTTELANNGLSTDSAME 120
 61 EMLNSTDTTPLTGPGTPESTVVEPARRSTGLDAGAVTELTTELANNGLSTDSAME 120
 DB 61 EMLNSTDTTPLTGPGTPESTVVEPARRSTGLDAGAVTELTTELANNGLSTDSAME 120
 QY 121 IOTTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTTPPATE 170
 121 IOTTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTTPPATE 170
 DB 121 IOTTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTTPPATE 180
 QY 171 AOTTOPGLEAQTAPAMEAQTAPAMEAQTTPPAMEAQTTPPAMEAQTAPATE 230
 171 AOTTOPGLEAQTAPAMEAQTAPAMEAQTTPPAMEAQTTPPAMEAQTAPATE 230
 DB 181 AOTTOPGLEAQTAPAMEAQTAPAMEAQTTPPAMEAQTTPPAMEAQTAPATE 240
 QY 231 AOTTOPATAQTTPPLAMEALSTEPSATEALSMEPTTKRGFLIPSVSVTHKGI 290
 231 AOTTOPATAQTTPPLAMEALSTEPSATEALSMEPTTKRGFLIPSVSVTHKGI 290
 DB 241 AOTTOPATAQTTPPLAMEALSTEPSATEALSMEPTTKRGFLIPSVSVTHKGI 300
 QY 291 SNLSVNPVGAPDHISVKOCLLAILLALVATTFVCTVLAVALSRKGMYPVRNYSPT 350
 291 SNLSVNPVGAPDHISVKOCLLAILLALVATTFVCTVLAVALSRKGMYPVRNYSPT 350
 DB 301 SNLSVNPVGAPDHISVKOCLLAILLALVATTFVCTVLAVALSRKGMYPVRNYSPT 360
 QY 351 EWCISSLLPDGGEGSATANGLSKAKSPGLTPREDREGDDLTHSFLP 402
 351 EWCISSLLPDGGEGSATANGLSKAKSPGLTPREDREGDDLTHSFLP 402
 DB 361 EWCISSLLPDGGEGSATANGLSKAKSPGLTPREDREGDDLTHSFLP 412

RESULT 13
 AAM56346
 ID AAM56346 standard; Protein; 412 AA.
 AC AAM56346;
 XX
 XX 27-JUL-1998 (first entry)
 DT
 XX
 XX P-selectin ligand protein.
 DE
 XX P-selectin ligand protein; glycoprotein; human; calcium-dependent;
 KW Interleukin adhesion; E-selectin; L-selectin; myocardial infarction;
 KW Inflammation; Crohn's disease; transplant rejection; rejection.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W09808949-A1.
 PN
 XX 05-MAR-1998.
 PD
 XX 29-AUG-1997; 97WC-US14159.
 PF
 XX 30-AUG-1996; 96US-0713556.
 PR
 XX (GEMV) GENETICS INST INC.
 PA
 XX Chang X, Cumming D, Kumar R, Larsen GR, Sako DS;
 XX Shaw G, Veldman GM;
 RI
 XX WPI; 1998-179434/16.
 DR
 XX N-PSDB; AAV22850.
 XX
 XX New isolated DNA encoding fusion protein including P-selectin ligand
 PT fragment - to direct second component, e.g. cytokine, to sites of
 PT selectin expression, used e.g. to stimulate bone and cartilage
 formation
 XX

PS Disclosure; Pages 68-69; 128p; English.
 XX
 CC The present sequence represents a P-selectin ligand protein. The
 CC P-selectin ligand is a glycoprotein. The protein binds in a
 CC calcium-dependent manner to P-selectin which is present on the surface
 CC of cells. The P-selectin ligand can be used to treat a wide variety of
 CC conditions characterised by intercellular adhesion involving P-, E- or
 CC L-selectins, e.g. myocardial infarction, infection, metastases,
 CC inflammation, Crohn's disease, and to prevent transplant rejection. It
 CC can also be used to raise specific antibody (useful therapeutically as
 CC inhibitors of adhesion or for immunodiagnosis of inflammation and cancer)
 CC or to screen for selective inhibitors. Fusion proteins of the P-selectin
 CC ligand protein are used to treat conditions related to selectin
 CC expression, e.g. fusion with interleukin-11 to treat bone marrow
 CC endothelial cells to stimulate megakaryocyte progenitors, bone
 CC morphogenic protein (BMP) to stimulate bone or cartilage formation,
 CC or other cytokines to sites of inflammation.
 CC
 XX SQ Sequence 412 AA;
 Query Match 99.3%; Score 2015; DB 19; Length 412;
 Best Local Similarity 97.6%; Pred. No. 3.4e-148;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEVLDYDFLPETEP 60
 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEVLDYDFLPETEP 60
 DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARDRQATEYEVLDYDFLPETEP 60
 QY 61 EMLNSTDTTPLTGPGTPESTVVEPARRSTGLDAGAVTELTTELANNGLSTDSAME 120
 61 EMLNSTDTTPLTGPGTPESTVVEPARRSTGLDAGAVTELTTELANNGLSTDSAME 120
 DB 61 EMLNSTDTTPLTGPGTPESTVVEPARRSTGLDAGAVTELTTELANNGLSTDSAME 120
 QY 121 IOTTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTTPPATE 170
 121 IOTTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTTPPATE 170
 DB 121 IOTTOPAA-----TEAOTPLAATEAOTRLTATEAOTPLAATEAOTTPPATE 180
 QY 171 AOTTOPGLEAQTAPAMEAQTAPAMEAQTTPPAMEAQTTPPAMEAQTAPATE 230
 171 AOTTOPGLEAQTAPAMEAQTAPAMEAQTTPPAMEAQTTPPAMEAQTAPATE 230
 DB 181 AOTTOPGLEAQTAPAMEAQTAPAMEAQTTPPAMEAQTTPPAMEAQTAPATE 240
 QY 231 AOTTOPATAQTTPPLAMEALSTEPSATEALSMEPTTKRGFLIPSVSVTHKGI 290
 231 AOTTOPATAQTTPPLAMEALSTEPSATEALSMEPTTKRGFLIPSVSVTHKGI 290
 DB 241 AOTTOPATAQTTPPLAMEALSTEPSATEALSMEPTTKRGFLIPSVSVTHKGI 300
 QY 291 SNLSVNPVGAPDHISVKOCLLAILLALVATTFVCTVLAVALSRKGMYPVRNYSPT 350
 291 SNLSVNPVGAPDHISVKOCLLAILLALVATTFVCTVLAVALSRKGMYPVRNYSPT 350
 DB 301 SNLSVNPVGAPDHISVKOCLLAILLALVATTFVCTVLAVALSRKGMYPVRNYSPT 360
 QY 351 EWCISSLLPDGGEGSATANGLSKAKSPGLTPREDREGDDLTHSFLP 402
 351 EWCISSLLPDGGEGSATANGLSKAKSPGLTPREDREGDDLTHSFLP 402
 DB 361 EWCISSLLPDGGEGSATANGLSKAKSPGLTPREDREGDDLTHSFLP 412

RESULT 14
 AAY29778
 ID AAY29778 standard; Protein; 412 AA.
 AC AAY29778;
 XX
 XX 04-NOV-1999 (first entry)
 DT
 XX
 XX Human P-selectin ligand protein SEQ ID NO:4.
 DE
 XX Human; P-selectin ligand; glycoprotein; fusion protein; infection;
 KW inflammation; intercellular adhesion; ulcerative colitis; asthma;
 KW diabetes; transplant rejection; myocardial infarction; thermal injury;
 KW metastatic condition; autoimmune thyroiditis; multiple sclerosis;
 KW Reynaud's syndrome; neutrophilic dermatitis; Sweet's syndrome;
 KW Grave's disease; glomerulonephritis; gingivitis; periodontitis;
 KW Crohn's disease; necrotising enterocolitis.
 XX
 XX Homo sapiens.
 OS

MN MO9943834-A2.
 XX
 PD 02-SEP-1999.
 XX
 PF 25-FEB-1999; 99WO-US04302.
 XX
 PR 27-FEB-1998; 98US-0032080.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Camphausen R, Chang X, Cumming D, Davis M, Kumar R;
 PI Larsen GR, Sako DS, Shaw G, Veldman GM;
 DR N-PSDB; AA208860.
 XX
 PT New P-selectin ligand fusion proteins, used for treating e.g.
 PT inflammation, infections, asthma, diabetes, ulcerative colitis or
 PT transplant rejection
 XX
 PS Disclosure; Page 110-111; 145pp; English.
 XX
 CC The present invention describes P-selectin ligand fusion proteins
 CC comprising amino acids 42-60, 42-402, 42-310, 42-88, 42-118 or 42-189
 CC of the P-selectin ligand protein. The fusion proteins comprising a
 CC P-selectin ligand act as ligands for P-selectin on human endothelial
 CC cells and platelets. The isolated P-selectin ligand proteins may be
 CC useful in treating conditions characterized by P-, E- or L-selectin
 CC mediated intercellular adhesion e.g. myocardial infarction, bacterial
 CC or viral infection, metastatic conditions, inflammatory disorders,
 CC thermal injury such as burns or frostbite, autoimmune thyroiditis,
 CC experimental allergic encephalomyelitis, multiple sclerosis, multiple
 CC organ injury syndrome secondary to trauma, diabetes, Reynaud's syndrome,
 CC neutrophilic dermatosis (Sweet's syndrome), inflammatory bowel disease,
 CC Grave's disease, glomerulonephritis, gingivitis, periodontitis,
 CC haemolytic uraemic syndrome, ulcerative colitis, Crohn's disease,
 CC necrotizing enterocolitis, granulocyte transfusion associated syndrome,
 CC or cytokine-induced toxicity. Isolated P-selectin ligand proteins may
 CC also be useful in organ transplantation, both to prepare organs for
 CC transplantation and to quell organ transplant rejection. P-selectin
 CC ligand proteins may be used to treat haemodialysis and leukophoresis
 CC patients or used as an antimetastatic agent. The fusion proteins can
 CC also be used to treat a condition which is affected by the protein to
 CC which the P-selectin ligand protein is fused. The fusion proteins can be
 CC used for the production of antibodies for use in therapy, detection,
 CC diagnosis and drug screening. The present sequence represents the human
 CC P-selectin ligand protein, as given in the present invention.
 CC
 SQ Sequence 412 AA;
 Query Match 99.3%; Score 2015; DB 20; Length 412;
 Best Local Similarity 97.6%; Pred. No. 3.4e-148;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 1 MPQLLLILLILGPGNSLQMDTWADAEKALGPLLARRRQATEYEVLDYDFLPETPP 60
 DB 1 MPQLLLILLILGPGNSLQMDTWADAEKALGPLLARRRQATEYEVLDYDFLPETPP 60
 QY 61 EMIRNSTDTPLTGPSTPSTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 DB 61 EMIRNSTDTPLTGPSTPSTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 QY 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 DB 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 QY 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 DB 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 QY 171 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 230
 DB 171 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 230
 QY 181 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 240
 DB 181 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 240
 QY 231 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 290
 DB 231 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 290
 QY 241 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 300
 DB 241 AOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 300

QY 291 SNLSVNVYGVADPHISVQCLIAIILIALVATIFPVCTVIAVRLSRKGMYPVRNYSPT 350
 DB 301 SNLSVNVYGVADPHISVQCLIAIILIALVATIFPVCTVIAVRLSRKGMYPVRNYSPT 360
 QY 351 EWCISLSLLPDGEGPSATANGSLSKAKSPGLTTPPREBREDDLTLSFLP 402
 DB 361 EWCISLSLLPDGEGPSATANGSLSKAKSPGLTTPPREBREDDLTLSFLP 412
 RESULT 15
 AA81967
 ID AA81967 standard; Protein; 412 AA.
 XX
 AC AA81967;
 XX
 DT 05-FEB-1999 (first entry)
 XX
 DE Human P selectin ligand glycoprotein #2.
 XX
 KW P selectin; ligand; glycoprotein; human; screening assay; inhibitor;
 KW E selectin; intercellular adhesion; inflammatory disease; treatment.
 XX
 OS Homo sapiens.
 XX
 PN US5843707-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 25-APR-1995; 95US-0428734.
 XX
 PR 25-APR-1995; 95US-0428734.
 PR 23-OCT-1992; 92US-0965662.
 PR 26-AUG-1993; 93US-0112608.
 PR 22-OCT-1993; 93WO-US10168.
 PR 28-APR-1994; 94US-0235398.
 PR 30-SEP-1994; 94US-0316305.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Chang X, Cumming D, Kumar R, Larsen GR, Sako DS;
 PI Shaw G, Veldman GM;
 DR N-PSDB; AA64998.
 XX
 PT DNA encoding P-selectin ligand protein - useful for producing
 PT recombinant protein
 XX
 PS Disclosure; Column 53-56; 67pp; English.
 XX
 CC This sequence represents a novel human P selectin ligand glycoprotein
 CC which can be used in screening assays for inhibitors of P- or
 CC E-selectin-mediated intercellular adhesion. The protein can be used
 CC in the treatment of inflammatory disease.
 SQ Sequence 412 AA;
 Query Match 99.3%; Score 2015; DB 20; Length 412;
 Best Local Similarity 97.6%; Pred. No. 3.4e-148;
 Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 1 MPQLLLILLILGPGNSLQMDTWADAEKALGPLLARRRQATEYEVLDYDFLPETPP 60
 DB 1 MPQLLLILLILGPGNSLQMDTWADAEKALGPLLARRRQATEYEVLDYDFLPETPP 60
 QY 61 EMIRNSTDTPLTGPSTPSTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 DB 61 EMIRNSTDTPLTGPSTPSTVEPARRSTGLDAGAVTELTTELANNGNLSTDSAAE 120
 QY 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 DB 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 QY 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170
 DB 121 IOTTGPA-----TEAOTPLAATEAOTTRLTATEAOTPLAATEAOTTPPAE 170

QY	171	AQTTQPTGLEAQTTAPAPAMEAQTTAPAPAMEAQTTAPAPAMEAQTTAPAPAMEAQTTAPAPATE	230
Db	181	AQTTQPTGLEAQTTAPAPAMEAQTTAPAPAMEAQTTAPAPAMEAQTTAPAPAMEAQTTAPATE	240
QY	231	AQTTQPTATAQTTPLAMEALSTEPSATEALSMETTKRGLFIPSVSVTHKGLPMAA	290
Db	241	AQTTQPTATAQTTPLAMEALSTEPSATEALSMETTKRGLFIPSVSVTHKGLPMAA	300
QY	291	SNLSVNPVGAPDHISVKOCLLAILLALVATTIFVCTVVLAVRLSRKGHMPVRNYSPT	350
Db	301	SNLSVNPVGAPDHISVKOCLLAILLALVATTIFVCTVVLAVRLSRKGHMPVRNYSPT	360
QY	351	EMVCISSLPDDGGEGPSATANGLSKAKSPGLTPEPREDRGDDLLHSLP	402
Db	361	EMVCISSLPDDGGEGPSATANGLSKAKSPGLTPEPREDRGDDLLHSLP	412

Search completed: July 14, 2003, 06:12:23
 Job time : 77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: July 14, 2003, 06:13:30 ; Search time 53 Seconds
(without alignments)
883.033 Million cell updates/sec

Title: US-09-825-580-2

Sequence: 1 MFLQLLLILLGPGNSIQ.....TPRPREDREGDDLTLSFLP 402

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 445758

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2030	100.0	402	US-10-163-853-1	Sequence 1, Appl
2	2030	100.0	402	US-10-211-786-2	Sequence 2, Appl
3	2030	100.0	402	US-09-819-157-1	Sequence 1, Appl
4	465	22.9	313	US-10-211-786-4	Sequence 4, Appl
5	209.5	10.3	688	US-09-864-761-36047	Sequence 36047, A
6	196.5	9.7	5179	US-10-025-380-1068	Sequence 1068, Ap
7	196.5	9.7	5179	US-09-922-217-1068	Sequence 1068, Ap
8	196.5	9.7	5179	US-09-833-263-1068	Sequence 1068, Ap
9	195	9.6	985	US-09-994-064-6	Sequence 6, Appl
10	195	9.6	985	US-09-994-064-6	Sequence 6, Appl
11	189.5	9.3	1031	US-09-815-243-10932	Sequence 10932, A
12	187	9.2	1367	US-09-801-368-10932	Sequence 108, App
13	183.5	9.0	331	US-09-864-761-36048	Sequence 36048, A
14	177.5	8.7	638	US-10-234-432-59	Sequence 59, Appl
15	176.5	8.7	638	US-10-121-988-143	Sequence 143, App
16	173	8.5	288	US-09-216-393-341	Sequence 341, App
17	173	8.5	288	US-09-216-393-344	Sequence 344, App
18	173	8.5	538	US-10-234-432-58	Sequence 58, Appl
19	173	8.5	727	US-10-234-432-57	Sequence 57, Appl

20	171.5	8.4	727	US-10-234-432-30	Sequence 30, Appl
21	171.5	8.4	5877	US-10-142-515-11	Sequence 11, Appl
22	171.5	8.4	5935	US-10-243-242A-8	Sequence 8, Appl
23	169.5	8.3	1387	US-10-156-761-13000	Sequence 13000, A
24	168	8.3	596	US-10-063-547-100	Sequence 100, App
25	168	8.3	596	US-10-174-590-310	Sequence 310, App
26	168	8.3	596	US-10-176-758-310	Sequence 310, App
27	168	8.3	596	US-10-063-616-100	Sequence 100, App
28	168	8.3	596	US-10-175-737-310	Sequence 310, App
29	168	8.3	596	US-10-063-502-100	Sequence 100, App
30	168	8.3	596	US-10-173-706-310	Sequence 310, App
31	168	8.3	596	US-10-173-738-310	Sequence 310, App
32	168	8.3	596	US-10-175-752-310	Sequence 310, App
33	168	8.3	596	US-10-176-482-310	Sequence 310, App
34	168	8.3	596	US-10-176-757-310	Sequence 310, App
35	168	8.3	596	US-10-176-913-310	Sequence 310, App
36	168	8.3	596	US-10-180-552-310	Sequence 310, App
37	168	8.3	596	US-10-180-557-310	Sequence 310, App
38	168	8.3	596	US-10-173-700-310	Sequence 310, App
39	168	8.3	596	US-10-174-572-310	Sequence 310, App
40	168	8.3	596	US-10-174-579-310	Sequence 310, App
41	168	8.3	596	US-10-174-582-310	Sequence 310, App
42	168	8.3	596	US-10-174-588-310	Sequence 310, App
43	168	8.3	596	US-10-175-739-310	Sequence 310, App
44	168	8.3	596	US-10-175-740-310	Sequence 310, App
45	168	8.3	596	US-10-175-743-310	Sequence 310, App

ALIGNMENTS

RESULT 1					
US-10-163-853-1					
Sequence 1, Application US/10163853					
Publication No. US20030050450A1					
GENERAL INFORMATION:					
APPLICANT: Cofitman, J.L., et al.					
TITLE OF INVENTION: Methods for Purifying Highly Anticlonic Proteins					
FILE REFERENCE: 01997.008800					
CURRENT APPLICATION NUMBER: US/10/163,853					
CURRENT FILING DATE: 2002-06-06					
PRIOR APPLICATION NUMBER: 60/296,402					
PRIOR FILING DATE: 2001-06-05					
NUMBER OF SEQ ID NOS: 1					
SOFTWARE: PatentIn version 3.1					
SEQ ID NO 1					
LENGTH: 402					
TYPE: PRT					
ORGANISM: homo sapiens					
US-10-163-853-1					
Query Match					
Best Local Similarity 100.0%; Score 2030; DB 9; Length 402;					
Matches 402; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	MFLQLLLILLGPGNSIQMDTMDAEKAGPLLRDRQATEYEDYFLPETEP	60		
DB	1	MFLQLLLILLGPGNSIQMDTMDAEKAGPLLRDRQATEYEDYFLPETEP	60		
QY	61	EWLRSTDTPLTGTPESTVEPARRSTGLDAGAVTELTTELANNGLSTDSAME	120		
DB	61	EWLRSTDTPLTGTPESTVEPARRSTGLDAGAVTELTTELANNGLSTDSAME	120		
QY	121	ITTPPAATEAQTPLAATEAQTTLVATEAQTPLAATEAQTTPPAATEOFTGIE	180		
DB	121	ITTPPAATEAQTPLAATEAQTTLVATEAQTPLAATEAQTTPPAATEOFTGIE	180		
QY	181	AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA	240		
DB	181	AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA	240		
QY	241	AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA	300		
DB	241	AQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA	300		

Db 241 AOTTPLAAMEALSTESPALEALNEPTTKGGLIFPSVSSVTHKGIIPMAASNLNVNYPVG 300
Qy 301 APDHISVKOCLAILILALVATIFPVCTVLAVALSRKGMYPVANSPTMWCISLILP 360
Db 301 APDHISVKOCLAILILALVATIFPVCTVLAVALSRKGMYPVANSPTMWCISLILP 360
Qy 361 DGEGGPSATANGLSKAKSPGLTPPEPREDEGDDLTLHSFLP 402
Db 361 DGEGGPSATANGLSKAKSPGLTPPEPREDEGDDLTLHSFLP 402

RESULT 2
US-10-211-786-2
; Sequence 2, Application US/10211786
; Publication No. US20030083258A1
; GENERAL INFORMATION:
; APPLICANT: Michael J. Eppihimer
; APPLICANT: Robert G. Schaub
; APPLICANT: Ronald Tuma
; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM
; FILE REFERENCE: 8702.0099-0000
; CURRENT APPLICATION NUMBER: US/10/211,786
; PRIOR FILING DATE: 2002-08-02
; PRIOR FILING DATE: 2001-08-03
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-786-2

Query Match 100.0%; Score 2030; DB 9; Length 402;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPQLQLLLILLIPGNSIQLMWTWADAEKALGPIIARDROATEYEYLDYDFLPETEP 60
Db 1 MPQLQLLLILLIPGNSIQLMWTWADAEKALGPIIARDROATEYEYLDYDFLPETEP 60
Qy 61 EMRNSTDTPLNGPGTPESTVEPAARSTGLDAGAVTELTELAMGNLSTDSAME 120
Db 61 EMRNSTDTPLNGPGTPESTVEPAARSTGLDAGAVTELTELAMGNLSTDSAME 120
Qy 121 IOTTPATEAOTTPLAATEAOTRLTATEAOTPLAATEAOTTPATEAOTTPG 180
Db 121 IOTTPATEAOTTPLAATEAOTRLTATEAOTPLAATEAOTTPATEAOTTPG 180
Qy 181 AOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATEAOTTPATE 240
Db 181 AOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATEAOTTPATE 240
Qy 241 AOTTPLAAMEALSTESPALEALNEPTTKGGLIFPSVSSVTHKGIIPMAASNLNVNYPVG 300
Db 241 AOTTPLAAMEALSTESPALEALNEPTTKGGLIFPSVSSVTHKGIIPMAASNLNVNYPVG 300
Qy 301 APDHISVKOCLAILILALVATIFPVCTVLAVALSRKGMYPVANSPTMWCISLILP 360
Db 301 APDHISVKOCLAILILALVATIFPVCTVLAVALSRKGMYPVANSPTMWCISLILP 360
Qy 361 DGEGGPSATANGLSKAKSPGLTPPEPREDEGDDLTLHSFLP 402
Db 361 DGEGGPSATANGLSKAKSPGLTPPEPREDEGDDLTLHSFLP 402

RESULT 3
US-09-819-157-1
; Sequence 1, Application US/09819157
; Patent No. US20020132991A1
; GENERAL INFORMATION:
; APPLICANT: Coffman, J.L., et al.
; TITLE OF INVENTION: METHODS FOR PURIFYING HIGHLY ANIONIC PROTEINS

; FILE REFERENCE: GFN-002
; CURRENT APPLICATION NUMBER: US/09/819,157
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/193,351
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 1
; Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-819-157-1

Query Match 100.0%; Score 2030; DB 10; Length 402;
Best Local Similarity 100.0%; Pred. No. 4e-137;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MPQLQLLLILLIPGNSIQLMWTWADAEKALGPIIARDROATEYEYLDYDFLPETEP 60
Db 1 MPQLQLLLILLIPGNSIQLMWTWADAEKALGPIIARDROATEYEYLDYDFLPETEP 60
Qy 61 EMRNSTDTPLNGPGTPESTVEPAARSTGLDAGAVTELTELAMGNLSTDSAME 120
Db 61 EMRNSTDTPLNGPGTPESTVEPAARSTGLDAGAVTELTELAMGNLSTDSAME 120
Qy 121 IOTTPATEAOTTPLAATEAOTRLTATEAOTPLAATEAOTTPATEAOTTPG 180
Db 121 IOTTPATEAOTTPLAATEAOTRLTATEAOTPLAATEAOTTPATEAOTTPG 180
Qy 181 AOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATEAOTTPATE 240
Db 181 AOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPAMEAOTTAPATEAOTTPATE 240
Qy 241 AOTTPLAAMEALSTESPALEALNEPTTKGGLIFPSVSSVTHKGIIPMAASNLNVNYPVG 300
Db 241 AOTTPLAAMEALSTESPALEALNEPTTKGGLIFPSVSSVTHKGIIPMAASNLNVNYPVG 300
Qy 301 APDHISVKOCLAILILALVATIFPVCTVLAVALSRKGMYPVANSPTMWCISLILP 360
Db 301 APDHISVKOCLAILILALVATIFPVCTVLAVALSRKGMYPVANSPTMWCISLILP 360
Qy 361 DGEGGPSATANGLSKAKSPGLTPPEPREDEGDDLTLHSFLP 402
Db 361 DGEGGPSATANGLSKAKSPGLTPPEPREDEGDDLTLHSFLP 402

RESULT 4
US-10-211-786-4
; Sequence 4, Application US/10211786
; Publication No. US20030083258A1
; GENERAL INFORMATION:
; APPLICANT: Michael J. Eppihimer
; APPLICANT: Robert G. Schaub
; APPLICANT: Ronald Tuma
; TITLE OF INVENTION: MODULATION OF LEUKOCYTE-ENDOTHELIAL INTERACTIONS FOLLOWING ISCHEM
; FILE REFERENCE: 8702.0099-0000
; CURRENT APPLICATION NUMBER: US/10/211,786
; PRIOR FILING DATE: 2002-08-02
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-786-4

Query Match 22.9%; Score 465; DB 9; Length 313;
Best Local Similarity 97.8%; Pred. No. 1e-25;
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MPQLQLLLILLIPGNSIQLMWTWADAEKALGPIIARDROATEYEYLDYDFLPETEP 60

Db 1 MPEQLLLILLGPGNSLQMDTMADEKALGPLLARDORQATEYEDYDPLPETEPP 60
Qy 61 EMLRNSTDTPPLTGPSTPESTTVEPARRST 91
61 EMLRNSTDTPPLTGPSTPESTTVEPARRPT 91
RESULT 5
US-09-864-761-36047
; Sequence 36047, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmlca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36047
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: MAP TO AB023048.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
US-09-864-761-36047
Query Match 10.3%; Score 209.5; DB 10; Length 688;
Best Local Similarity 24.3%; Pred. No. 4.2e-07;
Matches 66; Conservative 46; Mismatches 125; Indels 35; Gaps 5;
Qy 57 TEPPEMLRNSTDTPPLTGPSTPESTTVEPARRSTGLDCAVTEL-----TTELAM 109
Db 351 SETTTVSTTGTEIT-ITTSRGSEITTVTTAGSETTAVTTGSETTTTSGSETTVST 409
Qy 110 GNLSTDSAMEIQTTPAATEAQTTPPLATEAQTTLTATEAQ-----TPLAAT 159
Db 410 GSETTASPADLETTTSTSGSGTTTASGSETTVVITGSKTTTASRGSEATTVST 469
Qy 160 EAQTTPPAAT-----EAQTTPGTEAQTTPAAMEAQTTPAAMEAQTTPA 207
Db 470 SSETTASSTGSEMTVPFTVSEITTVSTIGSEATTSMAAGSEATTVSTGSETTASTA 529
Qy 208 AMEAQTTPAAMEAQTTPAATEAQTTPAATEAQTTPAAMEA-----LSTEPSATEAL 262
Db 530 GSETTASTAGSETTASTAGSETTACTGSETTSPSAGSETTATFTIGSESTIASTA 589
Qy 263 SMEPTTKRGLFIPFSVSVTHKGIPIAASNLS 294
Db 590 SLEPTATSLTSGSETTVSTITAGCATRATTVS 621
RESULT 6
US-10-025-380-1068
; Sequence 1068, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuguu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1068
Query Match 9.7%; Score 196.5; DB 9; Length 5179;
Best Local Similarity 24.5%; Pred. No. 3.9e-05;
Matches 89; Conservative 24; Mismatches 169; Indels 81; Gaps 13;
Qy 55 PETPEPEMLRNSTDTPPLTGPSTPESTTVEPARRSTGLDCAVTELTTTELAMNGNST 114
Db 3966 PTTTPTTTTTPPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 4024
Qy 115 DSAAMEIQTTPAATEA-OTTPAATEAQTTLTATEAQ--TTPLAATEAQTTPAATE 170

```

Db      4025 -----PTPRTGQTPTTPIITTTTIVTPRPRTGQTPTTPIITTTTIVTPRPRTG 40777
QY      171 AQTPTGLEAQT-----APAAEAQTAPAAEAQTTP-PAAMEAQTOTTAAEAQT 224
Db      4078 TOTPTTPIITTTTIVTPRPRTGQTPTTPIITTTTIVTPRPRTGQTPTTPIITTTT 4137
QY      225 -----APATEAQTQTPTATEAQTPLAAEALSTREBATEALSNPTTKGLFIPESV 279
Db      4138 VTPRPRTGQTPTTPIITTTTIVTP-----TPTPTGQTPTTPIITTTTIVTPRT 4190
QY      280 SVTHKGIPMAAS-----NLAVNYPVGAEDHISVKCCLAIILALVATIFEVCTVILAVR 334
Db      4191 TGTQTGPPHTNSTAPIALTLTNSP---PPESSTP----- 4222
QY      335 LSRGKMTPVNYSPTMVCISLLPDGGGPS-----ATANGLSKAKSPGLT 383
Db      4223 TSRS-----TSSLTESTILLSTLPAIEMTSTAPSTPTAPPTTSGGHTLPPPSST 4275
QY      384 PEP 386
Db      4276 TSP 4278

```

[illegible]

```
QY      280 SVTHKGIIPMAAS-----NLGVNPVGAARDHISVKQCCLALILIALVATIEPVCIVLAVR   334
        | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      4191 TGIQTGPPTHTSTAPIAEILTTSND--PPBSSTPQ-----                     4222
QY      335 LSRKGHHYPVRNYS-TEMVCISLLPDGGGPGS-----ATANGLSKAKSPGLT   383
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      4223 TSRS-----TSPLTESTITLTSLTPALEMTSTAPSTPTAFTTTSGGHITLSPDSTT  4275
QY      384 PEP    386
        |
Db      4276 TSP  4278
```

```

RESULT 8
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833.263
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-833-263-1068

```

Query Match	9.7%;	Score 196.5;	DB 10;	Length 5179;
Beat Local Similarity	24.5%;	Pred. No. 3.9e-05;		
Matches	89;	Conservative	24;	Mismatches 169; Indels 81; Gaps 13;

QY	55	PETEPRLNNSIDTTPLTGPTPESTTVPPAARSTGLDAGAVTELTTTELAMGNLSLT	114
DB	3966	PTTTTPTTTTTVTPPTPTGTGTPTTPTTTTTPPTPTGTGTPTTPTTTTTPVT	4024
QY	115	DSAMEIQTTPAATEA-QTTPLAATEAQTTLTATEAQ--TTPLAATEAQTPPAATE	170
DB	4025	-----PTPTGTGTQTPPTTPTTTTTPPTPTGTGTPTTPTTTTTPPTPTGTG	4077
QY	171	AQTTPGTGLEAQT-----APAMEAQTTAPAMEAQTP-PAAMEAQTTQTTMEAQT	224
DB	4078	TQTPPTTPTTTTTVTPPTPTGTGTPTTPTTTTTPPTPTGTGTPTTPTTTTT	4137
QY	225	-----APEATEAQTTQTPATEAQTTPLAMEALSTESATLEALSMETPKGLFI	279
DB	4138	VTPPTPTGTGTPTTPTTTTTPVT-----TPTPTGTGTPTTPTTTTTPPTPT	4190
QY	280	SVTHKGIIPMAAS-----NLSVNPVPGAPDHSVKQCLLALLLALVATIEFVCTVLAVR	334
DB	4191	TGQTGTPTHTNSTAPIALTLTNSN-PPSSSTQ-----	4222
QY	335	LSRKGHMYPVRNNSP-TEMVCISLLPDGGEGRS-----ATANGLSKAKSPGLT	383
DB	4223	TSTRS-----TSSPLTESTTLLSTLEPALEMTSTAPSTPAPTSTSGHTLSPPEST	4275
QY	384	PEP 386	
DB	4276	TSP 4278	

RESULT 9	US-09-994-064-6
;	Sequence 6, Application US/09994064
;	Publication No. US20030082788A1

GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/994,064
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/468,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-994-064-6

Query Match Best Local Similarity 9.6%; Score 195; DB 9; Length 985; Pred. No. 6.9e-06; Matches 93; Conservative 48; Mismatches 138; Indels 106; Gaps 16;

QY 14 PGNLSIQMDTWAD-----AEKALGPLAR-----DRQATEYEVLDYDF 53
DB 338 PGATLPFPDTAAPPDPTGSPPTTVPBPATITLIRSTSDMGFPSTABATGSETL----- 393
QY 54 LPETEPPEMLRNSTDTTPLTGPGTP--ESTVEPARARSTGLDAGAVTELTTTELANNGN 111
DB 394 ---SVVQETDRLSTTPLTLPLTPGESENTLFTTAPGISSTETPSAAHETTQTSAEIV 450
QY 112 LSTDSAAAEIQT-----TOPAATEAQT--TPLAATEAQTTLTATEAQT--T 154
DB 451 VFTQSPSTESSEARASQSOBPMWFTQTPTSEQAALTOQTQIAETEAFTQTPSAEQMFTQT 510
QY 155 PLAATEA-----QTPPAATEAQTQPTGLEAQTAPAM-----EAQTAPPA 197
DB 511 PAAETAPAQGTSTTPEITQTSKSTPPTARASAAPEVFTQSSSTVTEVFTQTPSTVAK 570
QY 198 AEAQTTPPA-----AMEAQTQTAM--EAQTAPPEAQT 234
DB 571 TLLSSTETPAIFRTQSAQTEAFTQSSAEPDTRQSTETHTFTQAPSTVPATQTPST 630
QY 235 QP-TATEAQT-PLAMEALSTEPSATEALSMEPT--TKRGLFIPFSVSVTHKGI PMA 290
DB 631 EPEVLQSPSTETVPFTRTLGAPEITQTPSAABEVYTRSSSTMPETAOST-----PLAS 685
QY 291 SNLSVN-----YVVGAPDH 304
DB 686 QNPSSGCTGHTNTEPRTTVPQTTPH 710

RESULT 10

US-09-994-064-66
Sequence 66, Application US/09994064
Publication No. US20030082788A1

GENERAL INFORMATION:

APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/994,064
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/468,190
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-994-064-66

Query Match Best Local Similarity 9.6%; Score 195; DB 9; Length 985; Pred. No. 6.9e-06; Matches 93; Conservative 48; Mismatches 138; Indels 106; Gaps 16;

QY 14 PGNLSIQMDTWAD-----AEKALGPLAR-----DRQATEYEVLDYDF 53
DB 338 PGATLPFPDTAAPPDPTGSPPTTVPBPATITLIRSTSDMGFPSTABATGSETL----- 393
QY 54 LPETEPPEMLRNSTDTTPLTGPGTP--ESTVEPARARSTGLDAGAVTELTTTELANNGN 111
DB 394 ---SVVQETDRLSTTPLTLPLTPGESENTLFTTAPGISSTETPSAAHETTQTSAEIV 450
QY 112 LSTDSAAAEIQT-----TOPAATEAQT--TPLAATEAQTTLTATEAQT--T 154
DB 451 VFTQSPSTESSEARASQSOBPMWFTQTPTSEQAALTOQTQIAETEAFTQTPSAEQMFTQT 510
QY 155 PLAATEA-----QTPPAATEAQTQPTGLEAQTAPAM-----EAQTAPPA 197
DB 511 PAAETAPAQGTSTTPEITQTSKSTPPTARASAAPEVFTQSSSTVTEVFTQTPSTVAK 570
QY 198 AEAQTTPPA-----AMEAQTQTAM--EAQTAPPEAQT 234
DB 571 TLLSSTETPAIFRTQSAQTEAFTQSSAEPDTRQSTETHTFTQAPSTVPATQTPST 630
QY 235 QP-TATEAQT-PLAMEALSTEPSATEALSMEPT--TKRGLFIPFSVSVTHKGI PMA 290
DB 631 EPEVLQSPSTETVPFTRTLGAPEITQTPSAABEVYTRSSSTMPETAOST-----PLAS 685
QY 291 SNLSVN-----YVVGAPDH 304

Db 686 QNPTSSGTGTHNTEPRTPVQTPH 710

RESULT 11

US-09-815-242-10932
Sequence 10932, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-10-23
PRIOR FILING DATE: 2000-10-23
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10932
LENGTH: 1031
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10932

Query Match 9.3%; Score 189.5; DB 10; Length 1031;
Best Local Similarity 22.4%; Pred. No. 1.8e-05;
Matches 64; Conservative 66; Mismatches 131; Indels 25; Gaps 8;

Db 3 LQLLLLLLGPNLSQLMD-TWADAEKALGPLLARDRROATEYELDYDLPEPTEPP 60
679 MQBALIDTTGGSSKQQLQLMWD--ELLEPLIGGFTPADQKRIHQ--LTQTIKT 734
QY 61 EML---RNSDTTPLPGTPESTYEPARRSTGLD-----AGCAVTELTTELAN 108
QY 735 ALLLVEKSTETTSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 794
Db 109 MGNLSTDSAMRIQTTPATEAQTTPATEAQTTPATEAQTTPATEAQTTPATEAQTTP 168
Db 795 TSSSTSPSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 854
QY 169 TEAQTTPGLEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPA 228
Db 855 TSSSTSPSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 914
QY 229 TEAQTTPATE-----AQTTPAAMEALSTESATBAISMETTK 269
Db 915 TSSSTSPSTLNENSOSKQNSVIVAVES--NODENDQSNS-KPSAK 958

RESULT 12

US-09-801-368-108
Sequence 108, Application US/09801368
Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT FILING DATE: 2001-03-07
PRIOR FILING DATE: 2001-03-07
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 2000-01-19
PRIOR FILING DATE: 1998-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 108
LENGTH: 1367
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

Query Match 9.2%; Score 187; DB 10; Length 1367;
Best Local Similarity 24.0%; Pred. No. 3.8e-05;
Matches 55; Conservative 36; Mismatches 110; Indels 28; Gaps 4;

QY 76 GTPSTYEPARRSTGLDAGCAVTELTTELAMGNLSTDSAMRIQTTPAATAQTTP 135
Db 207 GTKSSSTSSSTSSST-----TTSSSTSSSTSSSTSSSTTT 244
QY 136 LATEAQTTPATEAQTTPATEAQTTPATEAQTTPATEAQTTPATEAQTTPA 195
Db 245 STSSSTSSSTTAATPTTCTCKEKPPTTCTCKEKPPTTCTCKEKPPTTCTCK 303
QY 196 PAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAME 253
Db 304 KTCCTKTTTPVTPSSSTSSSAPVTPSSSTSSSAPVTPSSSTSSSAPVTPSS 363
QY 254 TTPSATLMEPTTKRGIFPVSYSVTHKGIPIAASLNVYVQAP 302
Db 364 TSSSAPVTS--STSSSAPVTSSTSSSAPVTPSSSTSSSAP 409

RESULT 13

US-09-864-761-36048
Sequence 36048, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR FILING DATE: 2000-02-04
PRIOR FILING DATE: 2000-02-04
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-08-03
PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36048
LENGTH: 331
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AB023048.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
US-09-864-761-36048

Query Match 9.0%; Score 183.5; DB 10; Length 331;
Best Local Similarity 25.8%; Pred. No. 1.2e-05;
Matches 61; Conservative 51; Mismatches 99; Indels 25; Gaps 6;
QY 57 TEPPELRNSTDTPLTGCTPESTVEPARSGTGLDGCAYT-----ELTTELAM 109
DB 53 SEPTMASTWSEETMASTIG-PEPTVSTASTASSVTVFAAGSETTASTVSGSTTVST 111
QY 110 GNSTSTSAAMEIOT-----TOPATEAOTPLAATE-AOTTLTATEAOTPLAAT 159
DB 112 GSETTASTASTGSTSTDSSTGSETTASTGSTSTTASTGSTASTTASTGSTTAST 171
QY 160 EAOTTPPAATEAOTPTGLEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAM 219
DB 172 DSETTTTCGSGEMTAVSTVFETTTASTGSTSTASTSTASTSTASTSTASTSTAST 231
QY 220 EAOTTPPAATEAOTPTGLEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAM 268
DB 232 GSETTKATVYTGSETTASTASTNGLETTTVFTIGSDTTASTGSTSTASTVASTGSEMT 287

RESULT 14
US-10-234-432-59
Sequence 59, Application US/10234432
Publication No. US20030091598A1

GENERAL INFORMATION:
APPLICANT: Homer, Mary J.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Pearing, David H.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: 210121.560
CURRENT APPLICATION NUMBER: US/10/234,432
CURRENT FILING DATE: 2002-08-30
NUMBER OF SEQ ID NOS: 108
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 59
LENGTH: 638
TYPE: PRT
ORGANISM: Babesia sp. WAI
US-10-234-432-59

Query Match 8.7%; Score 177.5; DB 9; Length 638;
Best Local Similarity 25.5%; Pred. No. 7.3e-05;
Matches 81; Conservative 24; Mismatches 110; Indels 103; Gaps 14;
QY 42 QATEYBYLDYDPLPETPEPEMLRNSTDTPLTGCTPESTVEPAA-----RSTGLDA 95
DB 42 QAT-----DVAPSSDQP---TODAQAPSPNP-TPASTVATPASPQSGSANQOQSOTGA 90
QY 96 GGAATVELTTELAMGNLSTDSAME-----IOTTPAA-- 128
DB 91 GESQPVLTSTWATVSEETVETPKTENNVVQATVTPAQPVAENVVSCPTQTVAPABA 150
QY 129 -----TEAQTPLA---ATEAQTTLTATEAQTPLAATEA----- 161
DB 151 PQQPAQVAPQATAGIQAQPOQVATETATAEQPVATTEVQ-WQAAAESAPILLETQ 209
QY 162 ---QTP-----PATEAQTTPGTEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 204
DB 210 VMTQTPAPEETQAPVTTESPAQOPQVAPQVAPQVAPQVAPQVAPQVAPQVAPQV 269
QY 205 PPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTPAAMEAQTTP 254
DB 270 QPQVQATQTPVAPQVAPQVAPQVAPQVAPQVAPQVAPQVAPQVAPQVAPQVAPQV 328
QY 255 ---EPSATEALSMEPTTK 269
DB 329 AQPQVAAEAQVQVQPVK 346

RESULT 15
US-10-121-988-143
Sequence 143, Application US/10121988
Publication No. US20030068327A1
GENERAL INFORMATION:
APPLICANT: Hosken, Nancy Ann
APPLICANT: McGowan, Patrick
APPLICANT: Sleath, Paul R.
APPLICANT: Mosaman, Sally P.
APPLICANT: Evans, Lawrence S.
APPLICANT: Swanson, Ryan M.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: 210121.538C1
CURRENT APPLICATION NUMBER: US/10/121,988
CURRENT FILING DATE: 2002-04-11
NUMBER OF SEQ ID NOS: 183
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 143
LENGTH: 699
TYPE: PRT
ORGANISM: HSV-2
US-10-121-988-143

Query Match 8.7%; Score 176.5; DB 9; Length 699;
Best Local Similarity 22.8%; Pred. No. 9.6e-05;
Matches 95; Conservative 49; Mismatches 167; Indels 105; Gaps 17;

QY 24 WA-----DEAKALGPIIARORRC-ATEYEYLDYDF-----LPETEPPEMLRNST 67
DB 300 WAPGALDDGYAPPPPPRRFRALRIDPEGVDPVRAPTGRRLMALTEDTSSDSPTSAP 359
QY 68 DTTPLTGPCTPESTVEPPAARSTGLDAGAVT---ELTTELAMGNLSTDSAMEIQTT 124
DB 360 EKTPLPVSATMMAFSVDPSAEP-----APATTTPDEMATAQATVAVTPETAV----AS 411
QY 125 QPAATEAQTTPLATEAQTT-----TRLVATEAQTTPLATEAQTTPPAATEAQTTOPTGL 179
DB 412 PPATASVSSPLPAAALATPGAGHTNTSSASAATP-PTTPAPTTPPTSTHATPRPTTP 470
QY 180 EAQTT-----APAAEAQTTAPAAEAQTTAPAAEAQTTQTAMEAQTTAPEATEAQ 232
DB 471 GPOTTTPGPATPGVGASAAAPTADSPITA--SPPATAPGPSAANVSVAATTATPGTRGTA 528
QY 233 TTOPT-----ATEAQTTPLAMEAL 252
DB 529 RTEPTDCKTPHGPADAPGSSPAPPPPEHKGPEEFEGAGDGEPPEDDSATGLAFTPN 588
QY 253 STEPSATEALSMETTRKGLFIPFSVSSVTHKGI PMAASNLVNYVPVG-APDHISVKQCL 311
DB 589 PNKPPPARPGPIRPTLPFGILGPLAPNTPR-----PPAQP- AKDMPSGPTPOHILPLFWFL 643
QY 312 LA-----LILALVATIFVCTVAVLRL-----SRKGMYPVRNYSPTENVCI 355
DB 644 TASPALDILFIISTTIHTAAFCVLAALQLMRGAGRRRYAHPSVRY-----VCL 694

Search completed: July 14, 2003, 06:22:20
Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model

Run on:

July 14, 2003, 06:11:00 ; Search time 26 Seconds
(without alignments)
454.924 Million cell updates/sec

Title: US-09-825-580-2

Perfect score: 2030
Sequence: 1 MFLQLLLLILLGPGNSIQL.....TPRPDRRGDDTLHSFLP 402Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2030	100.0	402	2	US-08-477-254A-2
2	2030	100.0	402	2	US-08-472-576B-2
3	2030	100.0	402	2	US-08-428-734B-2
4	2030	100.0	402	3	US-09-063-237-1
5	2030	100.0	402	4	US-08-713-556F-2
6	2015	99.3	412	2	US-08-472-576A-4
7	2015	99.3	412	2	US-08-472-576B-4
8	2015	99.3	412	2	US-08-428-734B-4
9	2015	99.3	412	4	US-08-713-556F-4
10	465	22.9	264	4	US-08-713-556F-42
11	465	22.9	313	4	US-08-713-556F-36
12	463	22.8	269	4	US-08-713-556F-38
13	278	13.7	437	4	US-08-713-556F-40
14	232	11.4	42	5	PCT-US96-10043-3
15	216	10.6	42	5	PCT-US96-10043-13
16	196.5	9.7	249	3	US-08-700-651-15
17	196.5	9.7	249	3	US-08-928-361B-20
18	195.5	9.6	1721	3	US-08-700-651-5
19	195.5	9.6	1721	3	US-08-928-361B-6
20	195	9.6	985	5	PCT-US96-03916-6
21	195	9.6	985	5	PCT-US96-03916-66
22	189.5	9.3	216	3	US-08-928-361B-27
23	187.5	9.2	786	4	US-09-103-429A-3
24	187.5	9.2	805	4	US-09-103-429A-3
25	182.5	9.0	216	3	US-08-928-361B-8
26	182.5	9.0	1837	3	US-08-928-361B-5
27	171.5	8.4	750	4	US-09-165-239A-4

28	167	8.2	878	4	US-09-556-706B-2	Sequence 2, App1
29	165.5	8.2	451	1	US-08-287-001A-2	Sequence 2, App1
30	165.5	8.2	451	5	PCT-US95-09941-2	Sequence 2, App1
31	158.5	7.8	1848	4	US-08-296-791-6	Sequence 6, App1
32	158.5	7.8	1848	5	PCT-US95-10661A-6	Sequence 6, App1
33	158	7.8	2476	2	US-08-276-967-2	Sequence 2, App1
34	157	7.7	1048	4	US-09-171-699-10	Sequence 10, App1
35	154.5	7.6	907	3	US-08-783-774-2	Sequence 2, App1
36	154.5	7.6	907	4	US-09-328-599A-1	Sequence 1, App1
37	154.5	7.6	907	5	PCT-US95-04611A-19	Sequence 19, App1
38	154	7.6	175	3	US-08-700-651-12	Sequence 12, App1
39	154	7.6	175	3	US-08-928-361B-17	Sequence 17, App1
40	153	7.5	941	4	US-07-757-022B-14	Sequence 14, App1
41	153	7.5	1022	4	US-07-757-022B-84	Sequence 84, App1
42	153	7.5	1038	4	US-07-757-022B-74	Sequence 74, App1
43	153	7.5	1049	4	US-07-757-022B-58	Sequence 58, App1
44	153	7.5	1140	4	US-07-757-022B-104	Sequence 104, App1
45	153	7.5	1270	4	US-07-757-022B-44	Sequence 44, App1

ALIGNMENTS

RESULT 1
US-08-477-254A-2
; Sequence 2, Application US/08477254A
; Patent No. 5827817
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Sako, Dianne
; APPLICANT: Chang, Xiao Jia
; APPLICANT: Veldman, Geertuida M.
; APPLICANT: Cumming, Dale
; APPLICANT: Kumar, Ravindra
; APPLICANT: Shaw, Gray
; TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEGAL AFFAIRS
; STREET: 87 CAMBRIDGE PARK DRIVE
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,254A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,662
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/112,608
; FILING DATE: 26-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10168
; FILING DATE: 22-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,398
; FILING DATE: 28-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,305
; FILING DATE: 30-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWN, SCOTT A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: GI 5213E-PCT
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-477-254A-2

Query Match 100.0%; Score 2030; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 1.9e-173;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLLLILLGPGNSQLMDTWADAEKALGPIIARDRROATEYEYLDYFLPETEPP 60
DB 1 MPQLLLILLGPGNSQLMDTWADAEKALGPIIARDRROATEYEYLDYFLPETEPP 60
QY 61 EMLRNSDTTPTLPGTPESTVEPARRSTGLDAGAVTELTELANNGNLSTDSAME 120
DB 61 EMLRNSDTTPTLPGTPESTVEPARRSTGLDAGAVTELTELANNGNLSTDSAME 120
QY 121 IOTTOPATEAOTTPAATEAOTTRILATEAOTTPLAATEAOTTPPAATEAOTTPGTE 180
DB 121 IOTTOPATEAOTTPAATEAOTTRILATEAOTTPLAATEAOTTPPAATEAOTTPGTE 180
QY 181 AOTTAPAMEAOTTPAPAMEAOTTPPAAMEAOTTTOTTAPEATEAOTTPATE 240
DB 181 AOTTAPAMEAOTTPAPAMEAOTTPPAAMEAOTTTOTTAPEATEAOTTPATE 240
QY 241 AOTTPLAAMEALSTESATLSEPTTKRGFLIPSVSVTHKGI PMAASNLVNVYPVG 300
DB 241 AOTTPLAAMEALSTESATLSEPTTKRGFLIPSVSVTHKGI PMAASNLVNVYPVG 300
QY 301 APDHISVKOCLAILIALVATIFVCTVVLAVRLSRKGMYPVNVSPTEWVCISLLP 360
DB 301 APDHISVKOCLAILIALVATIFVCTVVLAVRLSRKGMYPVNVSPTEWVCISLLP 360
QY 361 DGEGPSATANGLSKAKSPGLTPPEPREDEGDDLTHSFLP 402
DB 361 DGEGPSATANGLSKAKSPGLTPPEPREDEGDDLTHSFLP 402

RESULT 2

US-08-472-576B-2
Sequence 2, Application US/08472576B
Patent No. 5840679

GENERAL INFORMATION:

APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia
APPLICANT: Veldman, Geertuida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray
APPLICANT: Camphausen, Raymond
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDEPARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,576B
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213E-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-472-576B-2

Query Match 100.0%; Score 2030; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 1.9e-173;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLLLILLGPGNSQLMDTWADAEKALGPIIARDRROATEYEYLDYFLPETEPP 60
DB 1 MPQLLLILLGPGNSQLMDTWADAEKALGPIIARDRROATEYEYLDYFLPETEPP 60
QY 61 EMLRNSDTTPTLPGTPESTVEPARRSTGLDAGAVTELTELANNGNLSTDSAME 120
DB 61 EMLRNSDTTPTLPGTPESTVEPARRSTGLDAGAVTELTELANNGNLSTDSAME 120
QY 121 IOTTOPATEAOTTPAATEAOTTRILATEAOTTPLAATEAOTTPPAATEAOTTPGTE 180
DB 121 IOTTOPATEAOTTPAATEAOTTRILATEAOTTPLAATEAOTTPPAATEAOTTPGTE 180
QY 181 AOTTAPAMEAOTTPAPAMEAOTTPPAAMEAOTTTOTTAPEATEAOTTPATE 240
DB 181 AOTTAPAMEAOTTPAPAMEAOTTPPAAMEAOTTTOTTAPEATEAOTTPATE 240
QY 241 AOTTPLAAMEALSTESATLSEPTTKRGFLIPSVSVTHKGI PMAASNLVNVYPVG 300
DB 241 AOTTPLAAMEALSTESATLSEPTTKRGFLIPSVSVTHKGI PMAASNLVNVYPVG 300
QY 301 APDHISVKOCLAILIALVATIFVCTVVLAVRLSRKGMYPVNVSPTEWVCISLLP 360
DB 301 APDHISVKOCLAILIALVATIFVCTVVLAVRLSRKGMYPVNVSPTEWVCISLLP 360
QY 361 DGEGPSATANGLSKAKSPGLTPPEPREDEGDDLTHSFLP 402
DB 361 DGEGPSATANGLSKAKSPGLTPPEPREDEGDDLTHSFLP 402

RESULT 3

US-08-428-734B-2
Sequence 2, Application US/08428734B
Patent No. 5843707

GENERAL INFORMATION:

APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia
APPLICANT: Veldman, Geertuida M.

APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray
APPLICANT: Camphausen, Raymond
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,734B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213E-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-734B-2

Query Match 100.0%; Score 2030; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 1.9e-173;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARRRQATVEYLDYPLPETPP 60
1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARRRQATVEYLDYPLPETPP 60
61 EMLRNSDTTPTLPGTPESTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
61 EMLRNSDTTPTLPGTPESTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
121 IOTTPAATEAQTTPLAATEAQTTLTAATEAQTTPAATEAQTTPAATEAQTTPG 180
121 IOTTPAATEAQTTPLAATEAQTTLTAATEAQTTPAATEAQTTPAATEAQTTPG 180
181 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 240
181 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 240
241 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 300

241 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 300
301 APDHISVKQCLATLILALVATIFVYCTVVLVLRSRKQHWV PVRNYSPTMVCISLILP 360
301 APDHISVKQCLATLILALVATIFVYCTVVLVLRSRKQHWV PVRNYSPTMVCISLILP 360
361 DGEGPSATANGSLSKAKSPGLTPPEPRDREGDDITLHSLP 402
361 DGEGPSATANGSLSKAKSPGLTPPEPRDREGDDITLHSLP 402

RESULT 4
US-09-063-237-1
Sequence 1, Application US/09063237
Patent No. 6124267
GENERAL INFORMATION:
APPLICANT: McEver, Rodger P.
APPLICANT: Cumming, Richard D.
TITLE OF INVENTION: O-glycan inhibitors of Selectin Mediated
TITLE OF INVENTION: Inflammation Derived from PSGL-1
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30306-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/063,237
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/649,802
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF110CIP7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-063-237-1

Query Match 100.0%; Score 2030; DB 3; Length 402;
Best Local Similarity 100.0%; Pred. No. 1.9e-173;
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARRRQATVEYLDYPLPETPP 60
1 MPQLLLILLGPGNSLQMDTWADAEKALGPLARRRQATVEYLDYPLPETPP 60
61 EMLRNSDTTPTLPGTPESTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
61 EMLRNSDTTPTLPGTPESTVEPAARRSTGLDAGAVTELTTELANNGNLSTDSAA 120
121 IOTTPAATEAQTTPLAATEAQTTLTAATEAQTTPAATEAQTTPAATEAQTTPG 180
121 IOTTPAATEAQTTPLAATEAQTTLTAATEAQTTPAATEAQTTPAATEAQTTPG 180
181 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 240
181 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 240
241 AQTTPAAMEALSTEPSATEALSMEPTTKGLFIPSVSVTHKGI PMAASNLVNVY PVG 300

Qy	18	AQTTAPAAAEQAOTTAAPAAAEQAOTTPRAAAEAQOTTTMAEAQAAPAEKTAQAOTTPATE	240
Db	181	AQTTAPAAAEQAOTTAAPAAAEQAOTTPRAAAEAQOTTTMAEAQAAPAEKTAQAOTTPATE	240
Qy	241	AQTTPLAAAEALSTEPSATEALSMETPTKGLFIPFSVSATYHKGIIPMAASNLNVNTPVG	300
Db	241	AQTTPLAAAEALSTEPSATEALSMETPTKGLFIPFSVSATYHKGIIPMAASNLNVNTPVG	300
Qy	301	APRHISVYKQCLAILLILAVITTFPCTVYVLAVALSRKGHWYPRVANSPTFEMVCISLLP	360
Db	301	APRHISVYKQCLAILLILAVITTFPCTVYVLAVALSRKGHWYPRVANSPTFEMVCISLLP	360
Qy	361	DGEGPSATANGGLSRKAKSPGLTPREPREDEGDDLTLSFLP	402
Db	361	DGEGPSATANGGLSRKAKSPGLTPREPREDEGDDLTLSFLP	402

RESULT 5
 US-08-713-556F-2
 Sequence 2, Application US/08713556F
 Patent No. 6277975
 GENERAL INFORMATION:
 APPLICANT: Larsen, Glenn
 APPLICANT: Sako, Dianne
 APPLICANT: Chang, Xiao Jia
 APPLICANT: Veldman, Geertuida M.
 APPLICANT: Cumming, Dale
 APPLICANT: Kumar, Ravindra
 APPLICANT: Shaw, Gray
 TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LEGAL AFFAIRS
 STREET: 87 CAMBRIDGEPARK DRIVE
 CITY: CAMBRIDGE
 STATE: MA
 COUNTRY: USA
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/713,556F
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/965,662
 FILING DATE: 23-OCT-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/112,608
 FILING DATE: 26-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/10168
 FILING DATE: 22-OCT-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/235,398
 FILING DATE: 28-APR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/316,305
 FILING DATE: 30-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/428,734
 FILING DATE: 25-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: BROWN, SCOTT A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: GI 5213F
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-8224
 TELEFAX: (617) 876-5851

```

; INFORMATION FOR SEQ ID NO: 2:
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 402 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-08-713-556F-2

Query Match          100.0%;   Score 2030;   DB 4;   Length 402;
Best Local Similarity 100.0%;   Pred. No. 1,9e-173;
Matches 402;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0.

```

Qy	1	MPQLILLILLIPNSIQI	MDTWADAEKALGLIARDRORATEYEVLDYDL	PETERP	60
Db	1	MPQLILLILLIPNSIQI	MDTWADAEKALGLIARDRORATEYEVLDYDL	PETERP	60
Qy	61	EMLRNSTDTPLTGHGTPES	TTVEBPARRSTGLDAGAVTELTTELANMGNLSTDSAME	120	
Db	61	EMLRNSTDTPLTGHGTPES	TTVEBPARRSTGLDAGAVTELTTELANMGNLSTDSAME	120	
Qy	121	IOTTPPAATEAOTTPLATTEAOTRLRTTEAOTTPLATTEAOTT	PPAATEAOTTOPGLE	180	
Db	121	IOTTPPAATEAOTTPLATTEAOTRLRTTEAOTTPLATTEAOTT	PPAATEAOTTOPGLE	180	
Qy	181	AOTTAPAMEAOTTPPAMEAOTT	PPAMEAOTTQTTAMEAOTTAPATEAOTTOPATE	240	
Db	181	AOTTAPAMEAOTTPPAMEAOTT	PPAMEAOTTQTTAMEAOTTAPATEAOTTOPATE	240	
Qy	241	AOTTPLAAMEALSTEPSATEAL	SMEPTTKRGLPIFISVSSTYHKGIPMAASNLNVNYPVG	300	
Db	241	AOTTPLAAMEALSTEPSATEAL	SMEPTTKRGLPIFISVSSTYHKGIPMAASNLNVNYPVG	300	
Qy	301	APDHISVYKCLLAILLILAVATIP	VCYVLAVALSKGMYPYRNTSPTEMCISSLP	360	
Db	301	APDHISVYKCLLAILLILAVATIP	VCYVLAVALSKGMYPYRNTSPTEMCISSLP	360	
Qy	361	DGEGBSATANGLSKASKSPGL	TPREPRDRRGDDLTLHSFLP	402	
Db	361	DGEGBSATANGLSKASKSPGL	TPREPRDRRGDDLTLHSFLP	402	

RESULT 6
 US-08-477-254A-4
 Sequence 4, Application US/08477254A
 Patent No. 5827817
 GENERAL INFORMATION:
 APPLICANT: Larsen, Glenn
 APPLICANT: Sako, Dianne
 APPLICANT: Chang, Xiao Jia
 APPLICANT: Veldman, Geerttruida M.
 APPLICANT: Cumming, Dale
 APPLICANT: Kumar, Ravindra
 APPLICANT: Shaw, Gray
 TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LEGAL AFFAIRS
 STREET: 87 CAMBRIDGEPARK DRIVE
 CITY: CAMBRIDGE
 STATE: MA
 COUNTRY: USA
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,254A
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/965,662

FILED DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213E-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-254A-4

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 4.3e-172;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MFLQILLILILGPGNSIQIMDTWADAEKALGPIIARDROATEYELDYDFLPETEP 60
DB 1 MFLQILLILILGPGNSIQIMDTWADAEKALGPIIARDROATEYELDYDFLPETEP 60
QY 61 EMLRNSDTTPTLTGPTPESTVEPAARRSTGLDAGAVTELTTELANNGNSTDSAME 120
DB 61 EMLRNSDTTPTLTGPTPESTVEPAARRSTGLDAGAVTELTTELANNGNSTDSAME 120
QY 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTTPLAATEAOTTPPAATE 170
DB 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTTPLAATEAOTTPPAATE 180
QY 171 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 230
DB 181 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 240
QY 231 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 290
DB 241 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 300
QY 291 SNLSVNYPVGADPHISVKOCCLAILILALVATTFVCTVLAVALRSRKGMTPVRNYSPT 350
DB 301 SNLSVNYPVGADPHISVKOCCLAILILALVATTFVCTVLAVALRSRKGMTPVRNYSPT 360
QY 351 EMLVCSILLPDGEGSPSATANGLSKXSGRLTPREPREDEGDDTLHSLFLP 402
DB 361 EMLVCSILLPDGEGSPSATANGLSKXSGRLTPREPREDEGDDTLHSLFLP 412

RESULT 7
US-08-472-576B-4
Sequence 4, Application US/08472576B
Patent No. 5840679
GENERAL INFORMATION:
APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia
APPLICANT: Veideman, Geerttruida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray

APPLICANT: Camphausen, Raymond
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,576B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213E-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-576B-4

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 4.3e-172;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MFLQILLILILGPGNSIQIMDTWADAEKALGPIIARDROATEYELDYDFLPETEP 60
DB 1 MFLQILLILILGPGNSIQIMDTWADAEKALGPIIARDROATEYELDYDFLPETEP 60
QY 61 EMLRNSDTTPTLTGPTPESTVEPAARRSTGLDAGAVTELTTELANNGNSTDSAME 120
DB 61 EMLRNSDTTPTLTGPTPESTVEPAARRSTGLDAGAVTELTTELANNGNSTDSAME 120
QY 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTTPLAATEAOTTPPAATE 170
DB 121 IOTTPAA-----TEAOTPLAATEAOTTRLTATEAOTTPLAATEAOTTPPAATE 180
QY 171 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 230
DB 181 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 240
QY 231 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 290
DB 241 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 300

QY 291 SNLSVNPVGAPDHISVKQCLLAIIILAVATIFPVCTVVLAVRLSRKGMYPVRNYSPT 350
DB 301 SNLSVNPVGAPDHISVKQCLLAIIILAVATIFPVCTVVLAVRLSRKGMYPVRNYSPT 360
QY 351 EWCISLLPDDGEGSPATANGLSKXSPGLTPPEPREDEGDDLTLSFLP 402
DB 361 EWCISLLPDDGEGSPATANGLSKXSPGLTPPEPREDEGDDLTLSFLP 412

RESULT 8

US-08-428-734B-4
Sequence 4, Application US/08428734B
Patent No. 5843707
GENERAL INFORMATION:
APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia
APPLICANT: Veldman, Geertuida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray
APPLICANT: Camphausen, Raymond
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,734B
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213E-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-428-734B-4

Query Match 99.3%; Score 2015; DB 2; Length 412;
Best Local Similarity 97.6%; Pred. No. 4,3e-172;
Matches 402; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MFLQILLIILIGPGNSQLQMDTWADAEKALGPLARRRQATEXEYLDYFLPETEPP 60
DB 1 MFLQILLIILIGPGNSQLQMDTWADAEKALGPLARRRQATEXEYLDYFLPETEPP 60
QY 61 EMIRNSTDTPLTGPGTPESTVEPPAARSTGLDAGAVTELTTELANKGNSTDSAME 120
DB 61 EMIRNSTDTPLTGPGTPESTVEPPAARSTGLDAGAVTELTTELANKGNSTDSAME 120
QY 121 IOTTPA-----TEAOTTPAATEAOTTLTATEAOTTPAATEAOTTPAATE 170
DB 121 IOTTPAATEAOTTPAATEAOTTPAATEAOTTLTATEAOTTPAATEAOTTPAATE 180
QY 171 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 230
DB 181 AOTTPGTEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 240
QY 231 AOTTPATEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 290
DB 241 AOTTPATEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAAMEAOTTPAATE 300
QY 291 SNLSVNPVGAPDHISVKQCLLAIIILAVATIFPVCTVVLAVRLSRKGMYPVRNYSPT 350
DB 301 SNLSVNPVGAPDHISVKQCLLAIIILAVATIFPVCTVVLAVRLSRKGMYPVRNYSPT 360
QY 351 EWCISLLPDDGEGSPATANGLSKXSPGLTPPEPREDEGDDLTLSFLP 402
DB 361 EWCISLLPDDGEGSPATANGLSKXSPGLTPPEPREDEGDDLTLSFLP 412

RESULT 9

US-08-713-556F-4
Sequence 4, Application US/08713556F
Patent No. 6277975
GENERAL INFORMATION:
APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia
APPLICANT: Veldman, Geertuida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,556F
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:

STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,556F
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/428,734
FILING DATE: 25-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-713-556F-36

Query Match 22.9%; Score 465; DB 4; Length 313;
Best Local Similarity 97.8%; Pred. No. 8,7e-34;
Matches 89; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60
DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60

QY 61 EMLRNSTDTTPLTGPGTPESTVEPARRST 91
DB 61 EMLRNSTDTTPLTGPGTPESTVEPARRPT 91

RESULT 12
US-08-713-556F-38
Sequence 38, Application US/08713556F
Patent No. 6277975
GENERAL INFORMATION:
APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia
APPLICANT: Veldman, Geerttruida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray

TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,556F
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/428,734
FILING DATE: 25-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-713-556F-38

Query Match 22.8%; Score 463; DB 4; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.1e-33;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60
DB 1 MPQLLLILLGPGNSLQMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETEP 60

QY 61 EMLRNSTDTTPLTGPGTPESTVEPAR 88
DB 61 EMLRNSTDTTPLTGPGTPESTVEPAR 88

RESULT 13
US-08-713-556F-40
Sequence 40, Application US/08713556F
Patent No. 6277975
GENERAL INFORMATION:
APPLICANT: Larsen, Glenn
APPLICANT: Sako, Dianne
APPLICANT: Chang, Xiao Jia

APPLICANT: Veldman, Geertuida M.
APPLICANT: Cumming, Dale
APPLICANT: Kumar, Ravindra
APPLICANT: Shaw, Gray
TITLE OF INVENTION: NOVEL P-SELECTIN LIGAND PROTEIN
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CAMBRIDGE PARK DRIVE
CITY: CAMBRIDGE
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,556P
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,662
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/112,608
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10168
FILING DATE: 22-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,398
FILING DATE: 28-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,305
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/428,734
FILING DATE: 25-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWN, SCOTT A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI 5213P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-713-556F-40

Query Match 13.7%; Score 278; DB 4; Length 437;
Best local Similarity 92.9%; Pred. No. 7, 1e-17;
Matches 52; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 34 PLIARDRQATEYEVLDYDFLPETPEPEMLRNSTDTTPTLTGPTPESTVPEPAARR 89
DB 275 PLHKREKQATEYEVLDYDFLPETPEPEMLRNSTDTTPTLTGPTPESTVPEPAARR 330

RESULT 14
PCT-US96-10043-3
Sequence 3, Application PC/TUS9610043
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,213
FILING DATE: 14-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Iech, Karen F.
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 00786/284001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-10043-3

Query Match 11.4%; Score 232; DB 5; Length 42;
Best local Similarity 100.0%; Pred. No. 3, 3e-14;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QLMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETPEP 60
DB 1 QLMDTWADAEKALGPLLARDRQATEYEVLDYDFLPETPEP 42

RESULT 15
PCT-US96-10043-13
Sequence 13, Application PC/TUS9610043
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10043
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,213
FILING DATE: 14-JUN-1995

```

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 00786/284001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-10043-13

Query Match 10.6%; Score 216; DB 5; Length 42;
Best Local Similarity 100.0%; Pred. No. 8.9e-13;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 60 PEMLRNSTDTTPTLGGPSTPESTVEPARARSTGIDAGAVTE 101
Db 1 PEMLRNSTDTTPTLGGPSTPESTVEPARARSTGIDAGAVTE 42
```

Search completed: July 14, 2003, 06:16:03
Job time : 46 secs